DIGIROP Prediction models for severe retinopathy of prematurity

ALDINA PIVODIC

Department of Clinical Neuroscience
Institute of Neuroscience and Physiology
Sahlgrenska Academy, University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg, 2023

Cover illustration: *The Iris* by Aida Edvardsson

DIGIROP Prediction models for severe retinopathy of prematurity

aldina.pivodic@gu.se

ISBN 978-91-8069-191-8 (PRINT)
ISBN 978-91-8069-192-5 (PDF)
http://hdl.handle.net/2077/75181

Printed in Borås, Sweden 2023
Printed by Stema Specialtryck AB

To Philip, Edvin & Miro ❤️

TABLE OF CONTENTS

ABSTRACT

BACKGROUND: Retinopathy of prematurity (ROP), a preventable, potentially blinding eye disease, is primarily diagnosed in extremely preterm infants. Gestational age (GA) and birth weight (BW) are the most prominent risk factors. Routine ROP examinations are performed to identify the low proportion of infants who progress to needing treatment. In Sweden, ~30% of all screened infants are diagnosed with ROP, and 6% require treatment. Safe ROP prediction models can improve infant well-being and make screening efficient by identifying low- and high-risk infants.

AIM: The overall aim of the thesis was to develop and validate prediction models for severe ROP requiring treatment and propose a clinical decision support tool for safe and effective release of low-risk infants from ROP screening examinations. In addition, the natural course of the disease was described, and the prognostic value of the parenteral nutrition duration (PND) on ROP was demonstrated.

MATERIALS AND METHODS: The model development data originated from the Swedish national ROP register (SWEDROP). External validations included data from SWEDROP, Germany, and the US. *Paper I* included 6947 infants in the model development and 2122 in the external validation cohort. Corresponding figures for *Paper II* were 6991 and 1241, and for *Paper IV*, 8814 and 2325, respectively. *Paper III* included 1082 infants in its external validation. Extended Poisson models were used to develop DIGIROP-Birth requiring GA, BW, and sex in version 1.0, and PND ≥14 days in version 2.0 as input variables. Logistic regression models were used to develop DIGIROP-Screen, including the status and the age at the first ROP diagnosis besides DIGIROP-Birth risk estimates. GA-specific cut-offs were identified for the clinical decision support tool.

RESULTS: The instantaneous risk for ROP peaked around 12 weeks postnatal age, irrespective of GA at birth. Longer PND was strongly correlated to ROP severity, and faster progression. The risk for ROP differed for boys and girls over GA and PND. DIGIROP models released ~50% of infants from all ROP screening examinations and additionally ~25% during the screening process, while maintaining 100% sensitivity.

CONCLUSION: DIGIROP models may safely and efficiently release infants from unnecessary ROP examinations. The models appear superior to other currently available ROP models and are freely available as an online application (www.digirop.com).

KEYWORDS: preterm birth, retinopathy of prematurity, prediction models, screening, clinical decision support tool
ISBN 978-91-8069-191-8 (PRINT)
ISBN 978-91-8069-192-5 (PDF)

SAMMANFATTNING (SUMMARY IN SWEDISH)

Enligt de svenska riktlinjerna bör neonatal hjärt-lungräddning övervägas på barn från 22 gestationsveckors (GV) ålder och rekommenderas att utföras från 23 GV. Stora framgångar inom den neonatala intensivvården (NIV) samt >50% överlevnad bland de barn som har erbjudits NIV födda före 24 GV har lett till dessa rekommendationer. De förtidigt födda barnen riskerar på grund av sin prematuritet att drabbas av flera allvarliga sjukdomar. En av dem är prematuritetsretinopati (ROP), en neurovaskulär ögonsjukdom som vid svår progression och bristande behandling leder till blindhet. Framför allt drabbas barn födda före 28 GV, de extremt prematurfödda barnen. För att upptäcka ROP genomförs idag regelbundna rutinundersökningar världen över, i Sverige på alla barn födda före 30 GV. Dessa undersökningar kan vara påfrestande och påverka barnens välbefinnande.

Prediktionsmodeller kan användas som ett verktyg för att optimera screeningsprocessen. Syftet med denna avhandling har varit att utveckla och validera enkla och lättillgängliga prediktionsmodeller för behandlingskrävande ROP som på ett säkert sätt kan avskriva barnen från ”onödiga” rutinundersökningar, och därmed bidra till barnens välbefinnande, samt optimera resursanvändningen inom vården. Vidare skulle det naturliga sjukdomsförloppet, samt associationen mellan parenteral nutritionsdurationen (PND) och ROP beskrivas. Den första utvecklade modellen, DIGIROP-Birth, baseras endast på GV, födelsevikt och kön. Den andra utvecklade modellen, DIGIROP-Screen, lägger till information om ROP status och ålder vid första ROP diagnosen. Materialet som används härrör från det svenska nationella kvalitetsregistret för ROP (SWEDROP), och inkluderar ~7000 barn i den första utvecklingen av modellerna, *Arbete I* och *Arbete II*. I dessa arbeten ingår även en extern validering med hjälp av data från SWEDROP, Tyskland och USA. *Arbete III* är en extern validering på en nutida kohort från SWEDROP. *Arbete IV* uppdaterar DIGIROP-Birth med att byta ut standardiserad vikt mot vikt i gram, tillägg av ≥14 dagar PND som en proxyvariabel för allvarligt medicinskt tillstånd, samt uppdaterar parameterskattningar för DIGIROP-Screen baserat på de nya riskestimaten från DIGIROP-Birth. En extern validering ingår även i detta arbete på en nutida kohort från SWEDROP.

Den momentana risken för ROP uppvisade ett högsta värde kring 12 veckor postnatal ålder, oavsett GV vid födelse. Fler dagar med parenteral nutrition var starkt korrelerat till ROP, och snabbare progression av sjukdomen. Risken för ROP skilde sig mellan pojkar och flickor vid olika GV och PND. DIGIROP modellerna kunde avskriva ~50% av barnen från alla rutinundersökningar, och ytterligare ~25% under screeningsprocessen, samtidigt som de behöll hög känslighet. Modellerna verkar vara överlägsna andra, idag högt rankade, prediktionsmodeller för ROP.

DIGIROP modellerna finns fritt tillgängliga i en online applikation (www.digirop.com).

LIST OF PAPERS

This thesis is based on the following studies:

PAPER I **Pivodic A**, Hård AL, Löfqvist C, Smith LEH, Wu C, Bründer MC, Lagrèze WA, Stahl A, Holmström G, Albertsson-Wikland K, Johansson H, Nilsson S, Hellström A. Individual Risk Prediction for Sight-Threatening Retinopathy of Prematurity Using Birth Characteristics. *JAMA Ophthalmol. 2020 Jan 1;138(1):21-29.*

Referred to DIGIROP-Birth 1.0 in the text.

PAPER II **Pivodic A**, Johansson H, Smith LEH, Hård AL, Löfqvist C, Yoder BA, Hartnett ME, Wu C, Bründer MC, Lagrèze WA, Stahl A, Al-Hawasi A, Larsson E, Lundgren P, Gränse L, Sunnqvist B, Tornqvist K, Wallin A, Holmström G, Albertsson-Wikland K, Nilsson S, Hellström A. Development and validation of a new clinical decision support tool to optimize screening for retinopathy of prematurity. *Br J Ophthalmol. 2021 May 12:bjophthalmol-2020-318719.*

Referred to DIGIROP-Screen 1.0 in the text.

PAPER III **Pivodic A**, E H Smith L, Hård AL, Löfqvist C, Almeida AC, Al-Hawasi A, Larsson E, Lundgren P, Sunnqvist B, Tornqvist K, Wallin A, Holmstrom G, Gränse L. Validation of DIGIROP models and decision support tool for prediction of treatment for retinopathy of prematurity on a contemporary Swedish cohort. *Br J Ophthalmol. 2022 Mar 11:bjophthalmol-2021-320738.*

Referred to External validation in the text.

PAPER IV **Pivodic A**, Holmström G, Smith LEH, Hård AL, Löfqvist C, Al-Hawasi A, Larsson E, Lundgren P, Gränse L, Sunnqvist B, Tornqvist K, Wallin A, Johansson H, Albertsson-Wikland K, Nilsson S, Hellström A. Duration of Parenteral Nutrition and Risk for Retinopathy of Prematurity – Development and Validation of the Revised DIGIROP Clinical Decision Support Tool. *Accepted with some revisions.*

Referred to Updated DIGIROP models, DIGIROP-Birth 2.0, and DIGIROP-Screen 2.0 in the text.

RELATED PUBLICATIONS NOT INCLUDED IN THE THESIS

PAPER V **Pivodic A**, Johansson H, Smith LE, Löfqvist C, Albertsson-Wikland K, Nilsson S, Hellström A. Evaluation of the Retinopathy of Prematurity Activity Scale (ROP-ActS) in a randomised controlled trial aiming for prevention of severe ROP: a substudy of the Mega Donna Mega trial. *BMJ Open Ophthalmol. 2022 Apr 8;7(1):e000923.*

PAPER VI **Pivodic A**, Nilsson S, Stahl A, Smith LEH, Hellström A. Validation of the Retinopathy of Prematurity Activity Scale (ROP-ActS) using retrospective clinical data. *Acta Ophthalmol. 2021 Mar;99(2):201-206.*

PAPER VII **Pivodic A**, Hellström A. Correspondence to "Prediction of severe retinopathy of prematurity in 24-30 weeks gestation infants using birth characteristics". *J Perinatol. 2022 Mar;42(3):416-417.*

PAPER VIII Hellström A, Nilsson AK, Wackernagel D, **Pivodic A**, Vanpee M, Sjöbom U, Hellgren G, Hallberg B, Domellöf M, Klevebro S, Hellström W, Andersson M, Lund AM, Löfqvist C, Elfvin A, Sävman K, Hansen-Pupp I, Hård AL, Smith LEH, Ley D. Effect of Enteral Lipid Supplement on Severe Retinopathy of Prematurity: A Randomized Clinical Trial. *JAMA Pediatr. 2021 Apr 1;175(4):359-367.*

PAPER IX Hellström A, **Pivodic A**, Gränse L, Lundgren P, Sjöbom U, Nilsson AK, Söderling H, Hård AL, Smith LEH, Löfqvist CA. Association of Docosahexaenoic Acid and Arachidonic Acid Serum Levels With Retinopathy of Prematurity in Preterm Infants. *JAMA Netw Open. 2021 Oct 1;4(10):e2128771.*

ABBREVIATIONS

AA	Arachidonic acid
AAP	American academy of pediatrics
AI	Artificial intelligence
AIC	Akaike's information criterion
AUC	Area under the ROC curve
BPD	Bronchopulmonary dysplasia
BW	Birth weight
cdf	Cumulative distribution function
CI	Confidence interval
DHA	Docosahexaenoic acid
DSO	Days on supplemental oxygen
EPV	Events per variable
ESPGHAN	European society for pediatric gastroenterology, hepatology, and nutrition
ETROP	Early treatment for retinopathy of prematurity
FiO_2	Fraction of inspired oxygen
GA	Gestational age
G-ROP	Postnatal growth and retinopathy of prematurity
HR	Hazard ratio
ICROP	International classification of retinopathy of prematurity
IGF-1	Insulin-like growth factor 1
IGFBP3	Insulin-like growth factor binding protein 3
IVH	Intraventricular hemorrhage
LCPUFA	Long-chain polyunsaturated fatty acid
NEC	Necrotizing enterocolitis
NICU	Neonatal intensive care unit
NPV	Negative predictive value
OR	Odds ratio
PaO_2	Partial pressure of oxygen
PDA	Patent ductus arteriosus
pdf	Probability density function
PMA	Postmenstrual age
PNA	Postnatal age
PND	Parenteral nutrition duration
PPV	Positive predictive value
PROBAST	Prediction model study risk of bias assessment tool
PROGRESS	Prognosis research strategy
RW-ROP	Referral-warranted ROP
ROC	Receiver operating characteristic
ROP	Retinopathy of prematurity
SD	Standard deviation
SDS	Standard deviation score
SE	Standard error
SpO_2	Peripheral capillary oxygen saturation

SWEDROP	Swedish National Register for ROP
TRIPOD	Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis
TW-ROP	Treatment warranted ROP
VEGF	Vascular endothelial growth factor
WINROP	Weight, insulin-like growth factor-1, neonatal, ROP

1

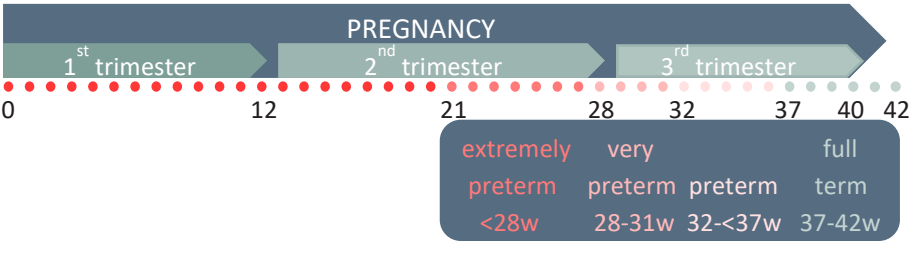
1 INTRODUCTION

1.1 FETAL DEVELOPMENT

The development from an egg to a full-grown fetus lasts ~38 weeks, starting at conception. The gestational age (GA) is calculated from the first day of the mother's last menstrual period.[1] Normal gestation lasts ~40 weeks. The pregnancy is divided into three trimesters, GA 0-12 weeks, GA 13-28, and 29-40 weeks.[2] The eyes develop early during the first trimester as extensions of the brain. The heart begins to rhythmically contract in week 6. By the end of the first trimester, all organs and the body structure are formed, nerves and muscles begin to work together, the retinal *vasculogenesis* (formation of new vessels) starts, and the embryo develops into a fetus that weighs ~20g and is ~8cm long. At week 16, the complete skeleton is built, the intestinal tract develops, and the skin starts to form. Retinal *angiogenesis* (formation of vessels from preexisting ones) starts in week 17-18. The fetus can hear and swallow in week 20. By week 22 retinal *vasculogenesis* is complete. In week 24, bone marrow begins to make blood cells, the lungs are formed, the fetus starts storing fat and weighs by this week ~700g and is ~30cm long. In week 32, the bones are fully formed, the eyes can open and close, lungs are not fully developed but start practicing breathing. By the end of the third trimester the brain and the neural system have developed, retinal *angiogenesis* is finalized, and the fetus grows.[2 3]

Multiple pregnancies and various complications, such as infections and chronic conditions, may lead to premature birth, defined as birth before 37 weeks of GA.[4] Birth between 28 and <32 weeks of GA is defined as very preterm and at <28 weeks of GA as extremely preterm.

FIGURE 1. PREGNANCY AND DEFINITION OF PREMATURITY. *w = weeks.*



1.2 PREMATURITY

EPIDEMIOLOGY

Yearly, one-tenth of all infants born worldwide are born prematurely, an increasing number that accounted for 15 million infants in 2010.[5,6] More than 60% of those are born in less developed regions, like in Africa and South Asia. About one million prematurely born infants die yearly due to preterm birth complications and an additional million due to other causes during the first month of postnatal life.[7] Twenty percent of preterm births occur very or extremely prematurely, at <32 weeks of GA.

Among ~110,000 infants currently yearly born in Sweden, ~5% are born preterm, ~1% before 32 weeks of GA, and ~0.3% are born extremely preterm before 28 weeks of GA, corresponding to ~350 infants.[8,9] From 2016, the Swedish guidelines recommend that neonatal cardiopulmonary resuscitation should be considered from 22 weeks GA and is recommended from 23 weeks GA, based on >50% survival during the first year of life in infants born at GA 22-23 weeks between 2014-2016 that received neonatal intensive care.[10,11] The *TINY* study, including all infants born at <24 weeks of GA 2007-2018 in Sweden, reported 62% live births, where 48% survived until 40 weeks of postmenstrual age (PMA).[12]

CAUSES

Multiple pregnancies, in vitro fertilization, previous premature birth, infections, chronic conditions such as hypertension and diabetes, unhealthy lifestyle, and stress are known causes of preterm birth.[13] However, many preterm births occur without any identified reason.

COMPLICATIONS

Infants born preterm are exposed to an increased risk of short- and long-term complications. Those include bleeding in the brain, i.e., intraventricular hemorrhage (IVH), hydrocephalus, heart problems such as an opening between

two important vessels called patent ductus arteriosus (PDA), respiratory problems requiring oxygen supplementation such as respiratory distress syndrome and bronchopulmonary dysplasia (BPD), problems with the digestive system such as necrotizing enterocolitis (NEC), hypothermia, anemia, hypoglycemia/hyperglycemia, sepsis, and other infections, and the eye disease retinopathy of prematurity (ROP).[12 14-16] In the Swedish *TINY* cohort including infants born before 24 weeks of GA, 51% had IVH, 17% severe IVH, 21% NEC, 90% had PDA, 18% persistent pulmonary hypertension, 91% had any ROP, 43% required ROP treatment, and 86% had BPD.[12 16]

Long-term disabilities include abnormal neurodevelopmental outcomes including cerebral palsy, epilepsy, cognitive impairment, vision impairment, hearing deficits, and impact on the pulmonary, renal, cardiovascular, and endocrine organ systems.[13 17-19]

NUTRITION

The *World Health Organization* (WHO), the *American Academy of Pediatrics* (AAP) and the *Committee of Nutrition of the European Society for Pediatric Gastroenterology, Hepatology and Nutrition* (ESPGHAN) recommend the mother's own milk with its unique composition to be the first choice of feeding in order to achieve optimal growth, development and immunological support for the infant.[20-22] Mothers to very preterm infants may, for different reasons, have lactation problems and difficulties providing breast milk to their children.[23] The alternative nutritional sources for preterm infants are donor human milk and preterm formula. Particularly very and extremely preterm infants are at high risk of being undernourished due to the sensitive and underdeveloped gastrointestinal tract resulting in limited uptake of the nutrients provided.[24 25] Therefore, parenteral nutrition is provided intravenously together with increasing amounts of enteral nutrition (through the gastrointestinal tract) based on the close monitoring of the infant's nutritional uptake and growth.[26 27] The duration of parenteral nutrition may be considered as a surrogate for the infant's morbidity status, since infants with critical illness require parenteral nutrition for a longer time period.[28] The optimal timing of initiation of the enteral and parenteral feeding, mode of feeding, duration and composition of the parenteral nutrition are debated. Positive impact on neurocognitive development was observed in

infants that had received recommended energy intake mainly enterally, and cautious approaches to aggressive parenteral nutrition intakes have been suggested.[29] Early start of enteral feeding is recommended as it may promote gut maturation, reduce feeding intolerance, accelerate achievement of full enteral feeding, decrease risk for growth restrictions, and late onset sepsis without increasing the incidence of NEC.[30]

GROWTH

Following preterm birth and disruption of the intrauterine supply of nutrients, very preterm infants lack sufficient levels of the insulin-like growth factor-1 (IGF-1), which are crucial for normal growth and development.[31 32] Hard et al. discussed associations of poor postnatal growth with intrauterine growth restriction, caused by increased metabolic rate precipitated by the adaptation to extrauterine life and further induced by severe diseases. The increased metabolic needs are not met. There is insufficient and/or non-optimal postnatal nutrition, and low levels of IGF-1.[33] Postnatal weight decreases immediately after birth are expected for all infants irrespective of GA, but prolonged return to BW is a potential risk factor for ROP.[34] Severe postnatal conditions, such as BPD, ROP and NEC, are associated with low BW and poor postnatal weight gain.[35]

To enable and improve infant growth monitoring references for weight, length and head circumference adjusted for GA and sex have been developed, such as Fenton's reference available from 22 weeks of GA.[36 37] The Swedish reference developed by Niklasson and Albertsson-Wikland based on 800,000 healthy singletons born during 1990-1999 is available from 24 weeks of GA.[38]

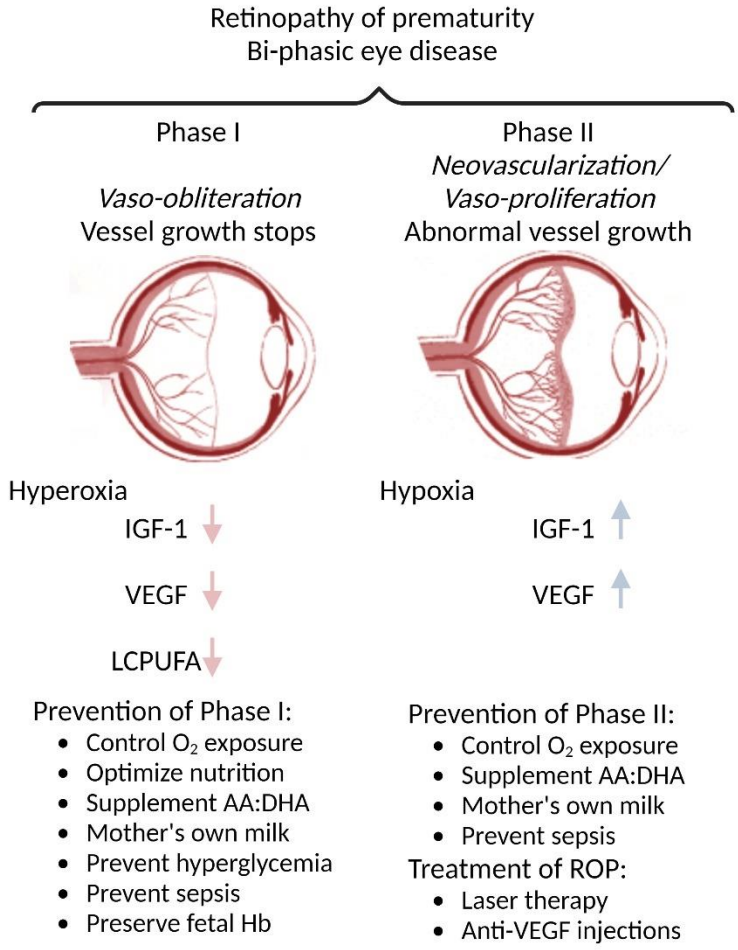
1.3 RETINOPATHY OF PREMATURITY

PATHOGENESIS

Retinopathy of prematurity, ROP, is a bi-phasic eye disease primarily diagnosed in extremely preterm infants. The end-stage of ROP was called *retrolental fibroplasia* in the early years when ROP was first time described by Terry in

1942.[39 40] The progression of ROP is characterized by cessation of vessel growth and *vaso-obliteration*, in the first phase, and is followed in severe cases by a second phase of abnormal vessel growth, *vaso-proliferation* or *neovascularization*.[3 40]

FIGURE 2. RETINAL VASCULARIZATION IN UTERO AND FOR PRETERM INFANTS DIAGNOSED WITH RETINOPATHY OF PREMATURITY. *IGF-1 = insulin-like growth factor-1, VEGF = vascular endothelial growth factor, LCPUFA = long-chain polyunsaturated fatty acids, ROP = retinopathy of prematurity, AA = arachidonic acid, DHA = docosahexaenoic acid, Hb = hemoglobin. Figure adapted from the original figure in Hellström et al 2013.[32] Created with BioRender.com.*



The first phase of ROP is initiated following preterm birth and there are likely many reasons for the cessation of vessel growth. One major cause is the increase in the partial pressure of oxygen (PaO_2) compared to that in utero, creating a relative *hyperoxia* that is further aggravated by supplemental oxygen provided to infants to keep them alive.[3] In this state *hypoxia*-driven vascular endothelial growth factor (VEGF) that promotes vascular growth, is downregulated.[41 42] Beside VEGF, normal *angiogenesis* requires attainment of certain levels of IGF-1, that is nutrient-dependent and suppressed following preterm birth but crucial for normal growth and development.[32 43 44] Additionally, loss of transfer of certain long-chain polyunsaturated fatty acids (LCPUFAs) from the mother to the fetus that normally occurs during the third trimester further exacerbates abnormal vascularization. The most important LCPUFAs for retinal development are ω-3 docosahexaenoic acid (DHA), ω-6 arachidonic acid (AA), and the ratio between the two.[42] In phase 2, the blood vessels grow abnormally initiated by the upregulated VEGF induced by the poor vascularization leading to retinal hypoxia, and by the increasingly higher IGF-1 levels in the maturing infant.[3 32 45]

EPIDEMIOLOGY

ROP is a leading cause of potentially preventable childhood blindness. In 2010, 19 million children were estimated living with ROP-related impaired vision worldwide, and 20.000 new cases of ROP-related blindness and/or severe visual impairment are reported yearly.[46] Three ROP epidemics were described, the first one in the 1940-1950s due to unrestricted oxygen support, the second and the third in 1960-1970 and 1990-2010 due to advances in neonatal intensive care in high- and middle-income countries, respectively, resulting in increased number of surviving preterm infants and larger ROP burden. As ROP screening programs improved over the years in high-income countries ROP-related childhood blindness decreased significantly.[47]

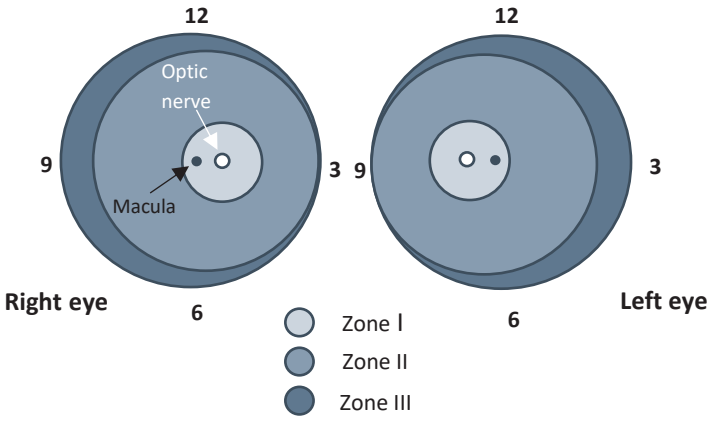
In Sweden, a national study including ~7000 routinely screened infants born at <31 weeks of GA and their ~40.000 examinations during the years 2008-2017, showed that 32% were diagnosed with any ROP.[48] Only 6% required treatment, decreasing with increased GA from 49% in infants born at <23 weeks of GA, to 0.3% in infants born at 29 weeks. No infants born at GA 30 weeks required treatment. Over the years, more immature infants were born, the incidence of

ROP remained similar, but the need for treatment increased. A population-based Swedish study identified 17 infants during years 2004-2017 who had ROP-related severe visual disability. Eleven (65%) of those could have been avoided if best practices were followed.[49]

CLASSIFICATION

The severity of ROP is classified by *stage* describing the appearance of the area between the vascular and avascular retina, *zone* describing the location of the disease, *extent* of the proliferative disease described by clock-hour designations, and the presence of *preplus/plus* disease describing the degrees of dilation and tortuosity of the central retinal vessels, as per the *International Classification of Retinopathy of Prematurity* (ICROP).[50-52] Stages 1-5 describe the acute phase of the disease, stage 1 characterized by a demarcation line between vascular and avascular area, stage 2 by a ridge between the two, and stage 3 by extraretinal neovascular proliferation or flat neovascularization. Stage 4 and 5 describe partial and total retinal detachment. Zone I is circular and the most central zone surrounding the optic nerve head. Zone II surrounds zone I and extends to the ora serrata nasally, and the area closest to zone I is referred to as posterior zone II. Zone III is present only temporally. *Regression*, *reactivation,* and *long-term sequelae* are described post-treatment. *Aggressive* ROP denotes a fast-progressing disease.

FIGURE 3. ZONE AND CLOCK HOURS USED TO DESCRIBE LOCATION AND EXTENT OF THE RETINOPATHY OF PREMATURITY. *Figure adapted from the original figure in Fierson 2018.*[51]



In 2019, ROP activity scale was developed and published by the members of the *International Neonatal Consortium* aimed to be implemented in clinical trials, following a request from the regulatory authorities.[53] This activity scale was based on severity classification suggested by nine experts in pediatric ophthalmology. The validation study, performed on a retrospective cohort, showed that the modification of the scale performed better than the conventionally used stage but not zone.[54] Using the data from a randomized controlled trial, the scale performed less well than the study's primary outcome, severe ROP (stage 3 or provided ROP treatment).[55]

Recently, a vascular severity score based on deep convolutional neural networks, a machine learning method, was proposed for monitoring of ROP post-treatment *regression* and *reactivation* with promising results.[56]

SCREENING

Based on a worldwide survey study performed by Mora et al. in 2017, 85% of the responding 92 countries had ROP screening programs. Non-respondents and those without any implemented ROP screening were mainly from Africa and former Soviet states.[57]

ROP screening guidelines are set country-wise and are dependent on the level of the neonatal intensive care and the incidence of complications in a specific country. In many high-income countries, like US, UK, Germany, and France, infants born at GA <31 weeks, in some also BW <1501 grams, are routinely screened for ROP.[51 58-60] Others screen infants that were more mature at birth.[61 62-64] Infants are screened as frequently as required based on the status and progression of the disease. Most infants' ROP screening is finalized by 45 weeks of PMA.

Since January 2020, all infants born before 30 weeks of gestation in Sweden are routinely screened for ROP.[48 65] Infants born at GA ≥30 weeks with severe medical conditions are screened at the discretion of their treating neonatologist. Between 2012 and 2020 all infants born at <31 weeks of gestation were to be screened, and before 2012 all infants with a GA <32 weeks were to be screened for ROP.[66 67] Repeated examinations are performed every other week to twice weekly depending on the ROP outcome.

Relevant data considering infant characteristics and disease status from the ROP screening examinations are reported into the *Swedish National Register for ROP* (SWEDROP).[68] As part of the *Swedish Neonatal Quality Register*, SWEDROP was initiated in 2007.[69] The register showed 97.6% coverage rate during years 2008-2017.[48] Perinatal data (GA, BW, sex, plurality, and comorbidities such as IVH, NEC, BPD), screening outcomes (date for first screening and number of screening examinations, first date for ROP stage 1-5, plus disease, maximum stage, most central zone, and extent (nasal and temporal), for left and right eyes), and information about ROP treatment (date and type of treatment provided at each treatment session) are registered.

DIAGNOSIS

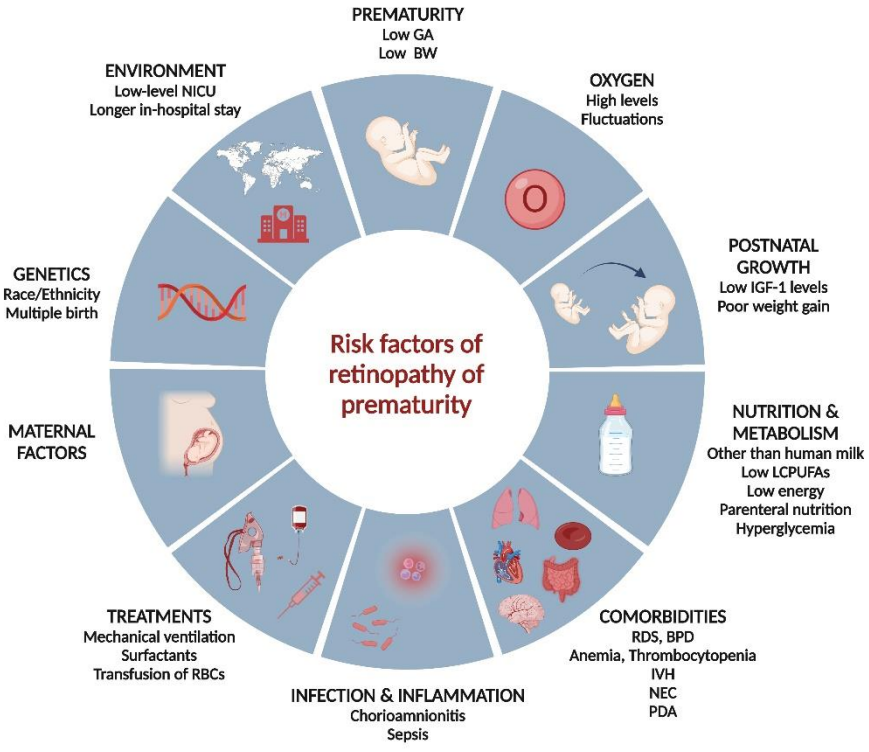
Diagnosis of ROP is made during the screening examinations by trained ophthalmologists using ophthalmoscopy, following ICROP.[50] Beside traditional ophthalmoscopy, objective photo documentation of the retinal fundus with techniques like RetCam (Pleasanton, CA, USA) are used, which is a contact fundus imaging system that is the most widely used system for tele-screening of ROP, showing high sensitivity and specificity.[70] Saved images facilitate monitoring of the disease over time, and enable obtaining second opinion of the diagnosis.

Owing to the increased utilization of the eye imaging systems, images as input data in the machine learning models have led to the development of screening-assisted diagnostic models in recent years. Deep convolutional neural networks and other machine learning models have been used to classify different zones, stages, plus disease, and aggressive ROP aimed to improve diagnosis.[71-78]

RISK FACTORS

The most prominent risk factors for ROP are low GA and low BW. Other important risk factors concern oxygen, postnatal growth, nutrition and metabolism, comorbidities, infections and inflammations, treatments, maternal factors, genetics, and environmental factors.[79]

FIGURE 4. RISK FACTORS OF RETINOPATHY OF PREMATURITY. *GA = gestational age, BW = birth weight, IGF-1 = insulin-like growth factor-1, LCPUFA = long-chain polyunsaturated fatty acids, RDS = respiratory distress syndrome, BPD = bronchopulmonary dysplasia, IVH = intraventricular hemorrhage, NEC = necrotizing enterocolitis, PDA = patent ductus arteriosus, RBC = red blood cells, NICU = neonatal intensive care unit. Created with BioRender.com.*



PARENTERAL NUTRITION AND RETINOPATHY OF PREMATURITY

The impact of parenteral nutrition on ROP as a factor associated with severe medical conditions, including infections, unhealthy gastrointestinal tract and abnormal microbiome, and as a potential risk factor vis-a-vis its improper nutritional composition, time of initiation and duration of treatment is yet to be explored in detail. Parenteral nutrition duration (PND) was first published as a risk factor for ROP in 1978 by Gunn et al. and was confirmed by Shohat et al. in 1983 in a cohort of 65 infants.[80 81] Vanhaesebrouck et al. studying a cohort of 412 infants born between 2000-2005 found that infants with ROP had twice the median days of PND than those without ROP.[82] Associations with ROP severity

and ROP treatment were not assessed in these publications. In a limited cohort of 118 infants, Niwald et al. showed that PND >10 days was a predictor for severe ROP requiring treatment.[83] In a cohort of 69 infants, Petrachkova et al. included PND >13 days in a prognostic model for type 1 ROP together with staying in the NICU >30 days, mechanical ventilation >30 days, RDS of 3rd degree and BPD.[84]

A higher total volume of parenteral nutrition has been found to be related to any ROP by Bassiouny, and to ROP treatment by Porcelli et al and Ali et al.[85 86 87]

PREVENTION

Preventative actions for ROP include control of oxygen exposure and fluctuations, optimization of nutrition, supplementation of LCPUFAs AA and DHA, provision of mother's own milk, preservation of fetal hemoglobin, and prevention of hyperglycemia, and sepsis.[33 79 88]

The Swedish *Mega Donna Mega* randomized controlled trial showed 50% reduction in severe ROP (stage 3 or severe ROP requiring treatment) in infants supplemented with AA:DHA up to 40 weeks of PMA compared to un-supplemented infants.[89] In a Norwegian study with a similar design, the AA:DHA group had a significantly decreased number of days with respiratory support but not severe ROP due to the low number of patients included, although a >50% numerical relative decrease was observed.[90] A secondary analysis of the *Mega Donna Mega* trial showed that increased circulating DHA for certain levels of AA decreased ROP severity, implying that the ratio beside the levels of the two LCPUFAs is important.[91]

TREATMENT

Type I ROP, defined as zone I any stage with plus disease, zone I stage 3 no plus disease, and zone II stage 2 or 3 with plus disease, should be treated according to the *Early Treatment for Retinopathy of Prematurity* (ETROP) criteria.[92 93] Treatment of ROP is performed by either laser therapy or anti-VEGF injections. There are several anti-VEGF therapies available on the market today, and the most commonly used in Europe and in the US are bevacizumab (*Avastin*), ranibizumab

(*Lucentis*), and aflibercept (*Eylea*). The use of these products raises concerns about the increased risk for *reactivation* of the disease and systemic effects of a single and repetitive treatments.[94] A meta-analysis shows that a single-treatment success rate is 89% with laser, 87% with bevacizumab, 81% with aflibercept and 74% for ranibizumab.[95]

1.4 PREDICTION MODELS

Prediction models are widely used in medicine. They might be used to help diagnose a disease by refining the inclusion criteria for routine screening of patients, to define the preventive intervention for a patient that has a high risk of developing a specific disease, or to inform the patient and the physician of the prognostic outcome of a disease. Generally, the *answer* we retrieve from a prediction model is a probability of a studied outcome of interest. By applying certain statistical functions on this probability with or without regard to relevant medical knowledge, a *qualitative answer* or *clinical decision* can be achieved that facilitates the application of a prediction model. Consequently, a *clinical decision support tool* is often the output of the model providing a decision recommendation.

Prediction modelling is divided into four parts, *development*, *validation*, *monitoring* including *update*, and *implementation*. Developing a prediction model requires careful consideration of various important factors including study design, sample size, selection of outcome of interest, selection of predictors, choice of statistical method, study of linear and non-linear relationships and interactions, overfitting, missing data, estimation, evaluation of the performance, visualization of the results and potential transfer to a clinical decision support tool for clinical usefulness. The final product is the model's public availability.[96] *Internal* validation and continuous *external* validations are expected to be performed to address the *generalizability* of the model on different populations and settings, and their *transportability* in time. Evaluation of the model's *discrimination*, *calibration*, and *clinical usefulness* should be performed.[97] Likewise, a monitoring plan should be put in place to evaluate the applicability of the model to future data including potential updates including *re-calibration*, *revision,* and *extension*.

1.5 PREDICTION MODELS FOR RETINOPATHY OF PREMATURITY

PUBLISHED MODELS

TRADITIONAL REGRESSION MODELS

The first published early ROP predicting model (WINROP) developed by Löfqvist et al. in Sweden in 2006 was based on GA, BW, and weekly weights, IGF-1 and IGFBP3 levels.[98] The second version of WINROP, published by Hellström et al. in 2009, was shown to function well without weekly IGF-1 and IGFBP3 levels.[99] Using its online application, the ROP monitoring tool WINROP2, is widely applied and has been evaluated on infants from different countries and settings.[100] The postnatal growth and retinopathy of prematurity (G-ROP) screening criteria was later developed by Binenbaum et al. on a cohort of ~7500 infants from the US and Canada, using BW, GA, weight gain 10-19, 20-29, and 30-39 days, and hydrocephalus status.[101]

TABLE 1. PREDICTION MODELS FOR RETINOPATHY OF PREMATURITY USING TRADITIONAL REGRESSION MODELS. *GA = gestational age, BW = birth weight, IGF-1 = insulin-like growth factor-1, IGFBP3 = insulin-like growth factor binding protein 3, IVH = intraventricular hemorrhage, NICU = neonatal intensive care unit, PNA = postnatal age, PMA = postmenstrual age, ROP = retinopathy of prematurity, RW-ROP = referral-warranted retinopathy of prematurity, SpO2 = peripheral capillary oxygen saturation, FiO2 = fraction of inspired oxygen. If more than one outcome was studied, sensitivity and specificity were presented for the more severe outcome.*

	No patients	Outcome	Predictors	Sensitivity Specificity
Including birth data only				
Yang and Donovan[102] 2009, US UHC	357	Pre- or threshold ROP	GA, BW, race, sex, multiple births	Sens 90% Spec 73%
Slidsborg et al[103] 2011, Denmark Danmark model 1	4182	ROP treatment	GA, BW	Sens 100% Spec 17%
Iu et al[104] 2022, China PW-ROP	1043	Type 1 ROP	GA, BW	Sens 87% Spec 79%
Including birth and postnatal data				

Hardy et al[105] 2003, US RM-ROP2	613	Pre-threshold ROP	GA, BW, race, hospital, multiple birth, onset/zone of ROP PMA, interval for ROP to pre-threshold, plus disease at first pre-threshold, stage 3	Not given
Löfqvist et al[98] 2006, Sweden WINROP1	79	ROP treatment	GA, BW, weekly weight, weekly IGF-1, IGFBP3	Sens 100% Spec 84%
Hellström et al[99] 2009, Sweden WINROP2	353	Stage 3	GA, BW, weekly weight gain	Sens 100% Spec 84%
Filho et al 2009, Brazil	317	Severe ROP	BW, IVH, weight gain proportion	Sens 66% Spec 63%
Binenbaum et al[106] 2011, US & Canada PINT-ROP	367	Severe ROP	GA, BW, daily weight gain rate	Sens 99% Spec 30%
Binenbaum et al[107] 2012, US & Canada CHOP-ROP	524	Type 1 or 2 ROP Type 1 ROP	GA, BW, weight gain rate (daily and weekly)	Sens 100% Spec 51%
Eckert et al[108] 2012, Brazil ROPScore	474	Any ROP ROP treatment	GA, BW, weight gain, oxygen, and blood transfusion up to 6 weeks PNA	Sens 96% Spec 56%
Ying et al[109] 2015, US e-ROP	979	RW-ROP	GA, BW, sex, race, multiple births, pre-plus quadrants, ROP stage, retinal hemorrhage, respiratory status, weight gain to first study exam	Sens 96% Spec 53%
Cao et al[110] 2016, US CO-ROP	858	Any ROP Type 1 or 2	GA, BW, weight gain up to 28 days of PNA	Sens 98% Spec 24%
Slidsborg et al[111] 2016, Denmark Denmark model 2	6490	ROP treatment	GA, SGA, sex, multiple birth, mechanical ventilation, blood transfusion	Sens 100% Spec 37%
Ricard et al[112] 2017, US & Germany STEP-ROP	627	Any ROP Type 1 or 2	GA, BW, respiratory distress syndrome, non-hispanic mother, multiple births	Sens 95% Spec 17%
Ludwig et al[113] 2017, US SUNDROP	843	TW- ROP	GA, BW, weight gain rate	Not given
Binenbaum et al[114] 2017, US CHOP-ROP update	524	Type 1 or 2 ROP Type 1 ROP	GA, BW, weight gain rate	Sens 100% Spec 11%
Gurwin et al[115] 2017, US TARP	242	Type 1 ROP	CHOP-ROP + e-ROP	Sens 100% Spec not given
Binenbaum et al[101] 2018, US & Canada G-ROP	7483	Type 1 or 2 ROP ROP treatment	GA, BW, daily weight gain, hydrocephalus	Sens 100% Spec 30%
McCauley et al[116] 2018, US OMA-ROP	191	Type 1 ROP	GA, BW, weight gain rate	Sens 100% Spec 62%
Ahmed & Badeeb[117] 2019, Egypt Alex-ROP	560	Any ROP Type 1 or 2	GA, BW, weight gain ratio at 7, 14, 21, 28	Sens 100% Spec 50%

Ying et al[118] 2019, US	7483	Any ROP Type 1 or 2	GA, BW, Apgar score 1min, maternal race, birth location and intubation	Not given
Petrachkova et al[84] 2019, Russia	69	Type 1 ROP	Staying in NICU >30 days, staying on mechanical ventilator >30 days, parenteral nutrition >13 days, respiratory distress syndrome of 3rd degree, BPD	Sens 88% Spec 69%
Cheng et al[119] 2019, US e-ROP	1239	Plus disease	GA, race, respiratory support, postnatal weight gain, images on ROP stage preplus/plus and blot hemorrhage	Sens 94% Spec 81%
Estrada et al[120] 2022, US & Canada G-ROP DSO	8949	Type 1 or 2 ROP ROP treatment	G-ROP + days on supplemental oxygen	Sens 100% Spec 23%
Siswanto et al[121] 2023, Indonesia FiO2 SpO2	160	Type 1 ROP	GA, Oxygen supplementation, socioeconomic status, outborn infant, lowest SpO2	Sens 100% Spec 73%

MACHINE-LEARNING MODELS

Unlike the development of diagnostic models to date, few prognostic ROP models have been developed using machine learning techniques. Coyner et al. developed in 2021 a model using only GA and one retinal fundus image at PMA 32-33 weeks with high accuracy.[122] Similar results were obtained by Wu et al. using longitudinal images and 46 clinical variables.[123]

TABLE 2. PREDICTION MODELS FOR RETINOPATHY OF PREMATURITY USING MACHINE-LEARNING TECHNIQUES. *GA = gestational age, ROP = retinopathy of prematurity, VSS = vascular severity scale.*

	No eyes/infants	**Outcome**	**Predictors**	**Sensitivity Specificity**
Using images with birth and postnatal data				
Coyner et al[122] 2021, USA	1579 eyes	ROP treatment	GA, VSS from one retinal fundus image at 32-22 weeks of PMA	Sens 100% Spec 49%
Wu et al[123] 2022, China	815 infants	Any ROP ROP severity	46 clinical variables and RetCam images	Sens 100% Spec 47%

1.6 SURVIVAL MODELS

CONCEPT

Time-to-event or survival analysis is often applied in clinical and epidemiological studies where time to an event is of importance. The key concepts of time-to-event analysis are *hazard* (instantaneous rate) and *survival* (probability to be free from the event) functions, described in detail below based on Harrell F.E. and Rodriguez G.[124 125] To exemplify, the hazard function of mortality in a general population described over ages is high near birth, accounting for an increased risk of complications leading to death at birth. After birth, it drastically decreases to low hazard. Then, it is stable for younger ages and increases for higher ages. The survival function is a probability to remain free from an event up to a specific time point. It is a probability with the value of 1.0 at time 0 (100% free from the event) which decreases over time. The survival function is called *survival function* also for events other than mortality.
Following questions may be studied by applying these methods:

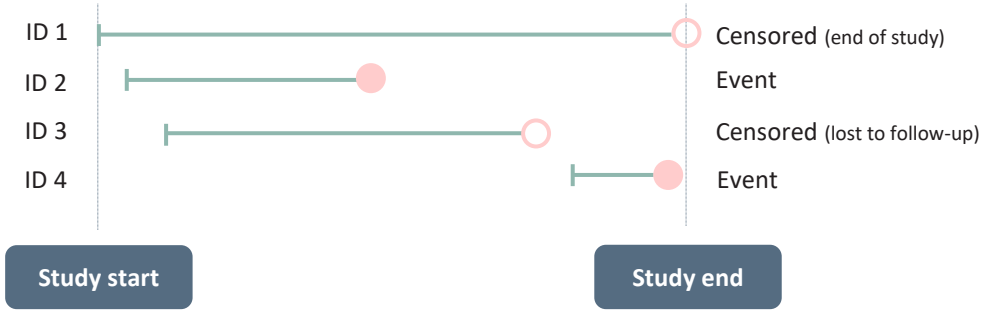
- *Describe the time course over PNA and PMA of the instantaneous rate (hazard function) of severe ROP requiring treatment in ROP-screened infants.*
- *Assess the difference in the effect of a preventive AA:DHA supplementation compared to conventional nutrition up to 40 weeks of postmenstrual age on severe ROP in a randomized controlled trial.*
- *Study strength and shape of prognostic factors for time to death in prematurely born infants in an observational epidemiological study.*

DATA

The simplest data set structured to perform survival analysis needs to have the following variables defined for each participant: the *time to event or censoring*, calculated as the difference between the baseline date and the date of the event or the last follow-up date if the event has not occurred, also known as *right-censoring* date, and the *status/indicator of event* (yes or no). Participants may enter the study at different time points and have different duration of the study follow-up. In studies where all participants are followed for an equal time, a more simple analysis may be utilized, such as logistic regression, unless the hazard function or

survival function of the disease are of specific interest to be described. Likewise, in the studies where all participants have reached the studied event during the follow-up, regression models may be used studying time-to-event as dependent variable.

FIGURE 5. EXAMPLE DATA FOR SURVIVAL ANALYSIS. *Time to event or censoring is the length of the green line.*



Other important nomenclature not further discussed within this thesis, nor applied in the project, are *interval-censoring*, *left-truncation*, *left-censoring*, *delayed entry*, and *competing risks*.

PDF, CDF, SURVIVAL, HAZARD, CUMULATIVE HAZARD, LOG-LIKELIHOOD

Assume that T is a non-negative continuous random variable representing time until a certain event of interest. Its probability density function (pdf) is denoted by $f(t)$ and its cumulative distribution function (cdf) $F(t) = Pr\{T < t\}$. $F(t)$ is the probability that an event has occurred by time t and is related to f by

$$F(t) = Pr\{T < t\} = \int_0^t f(x)dx$$

i.e., derivative of $F(t)$ is $f(t)$

$$F'(t) = f(t)$$

The complement of the cdf is the *survival* function, the probability that an event has not occurred during time t, expressed by

$$S(t) = Pr\{T \geq t\} = 1 - F(t)$$

i.e., the derivative of $S(t)$ is $-f(t)$

$$S'(\mathrm{t}) = -f(t)$$

The *hazard* function is defined by

$$\lambda(t) = \lim_{dt \to 0} \frac{Pr\{t \leq T < t + dt \mid T \geq t\}}{dt}$$

The numerator represents the conditional probability that an event occurs in the interval $[t, t + dt]$. The denominator is the width of the interval. By taking the ratio and the limit that approaches *zero* we get the instantaneous rate of the event.

Using the law of conditional probability, the expression above may be re-written as

$$\lambda(t) = \lim_{dt \to 0} \frac{Pr\{t \leq T < t + dt\}/dt}{\Pr\{T \geq t\}}$$

$$= \lim_{dt \to 0} \frac{[F(t + dt) - F(t)]/dt}{S(t)}$$

$$= \frac{\frac{d}{dt}F(t)}{S(t)}$$

$$= \frac{f(t)}{S(t)}$$

interpreted as the event rate at duration t.

Applying the derivative of $S(t)$, $-f(t)$, and derivation rules of the logarithm of a function, the expression may be re-written to

$$\lambda(t) = \frac{-\frac{d}{dt}S(t)}{S(t)} = -\frac{d}{dt}logS(t)$$

Integrating from 0 to t on both sides we get following expression for the survival function

$$S(t) = exp\left(-\int_0^t \lambda\,(x)dx\right)$$

The integral in the expression above is called *cumulative hazard*

$$\Lambda(t) = \int_0^t \lambda\,(x)dx$$

The other functions may be obtained given any of $f(t)$, $F(t)$, $S(t)$, $\lambda(t)$ and $\Lambda(t)$. Time-to-event data is most commonly studied by modelling $\lambda(t)$, known as hazard rate models, such as Cox regression models, further described below.[126] Others, like Royston and Parmar, model log $\Lambda(t)$, not discussed further in this work.[127]

The model parameters are solved through maximizing the likelihood function L, constructed using each individual's contribution to the statistical problem, using maximum likelihood estimation. For practical reasons log-likelihood $logL$ is maximized rather than likelihood L itslef.

Assume that we have n individuals included in a study, and that individual i is followed up until time t_i. If the individual gets event at t_i then the contribution to the likelihood function is the density at that duration

$$L_i = f(t_i) = S(t_i)\lambda(t_i)$$

If the individual is still free from event at t_i, i.e. censored at t_i, the contribution to the likelihood function is

$$L_i = S(t_i)$$

Considering all n individuals the likelihood may be expressed by

$$L = \prod_{i=1}^{n} L_i = \prod_{i=1}^{n} \lambda\,(t_i)^{d_i} S(t_i)$$

where d_i is the event indicator.

The log-likelihood (natural logarithm) is then expressed by

$$logL = \sum_{i=1}^{n} \left(d_i log\lambda(t_i) - \Lambda(t_i)\right)$$

HAZARD RATE MODELS

NON-PARAMETRIC PROPORTIONAL HAZARDS MODEL

Introducing a vector of explanatory variables or predictors $\boldsymbol{x}$ that are related to the survival time T, Cox proposed a family of hazards models in 1972, with the *proportional hazards model* being the most simple and most familiar.[126]

$$\lambda_i(t|x_i) = \lambda_0(t)exp(\boldsymbol{\beta x_i})$$

where $\lambda_0(t)$ is the baseline hazard function describing the risk for participants having $x_i = 0$, and $exp(\boldsymbol{\beta x}_i)$ is additional relative risk defined by the predictors $\boldsymbol{x}$, where $\boldsymbol{\beta x}_i = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$ for k predictors, with $exp(\beta_k)$ being known as hazard ratio (HR) for predictor k.

Taking the log of the hazard function we get an additive model, assuming proportional hazard owing to the $\lambda_0(t)$ not being dependent on X, i.e. the effect of X is same for all t.

$$log\lambda_i(t|x_i) = log\lambda_0(t) + \beta x_i$$

The cumulative hazard function is then

$$\Lambda_i(t|x_i) = \Lambda_0(t)exp(\beta x_i)$$

and the survival function

$$S_i(t|x_i) = S_0(t)^{exp(\beta x_i)}$$

The baseline hazard is in the proposal from Cox left completely unspecified and estimated non-parametrically. The model focuses on estimating β coefficients.

PARAMETRIC PROPORTIONAL HAZARDS MODEL

The parametric proportional hazards models have a specific functional form of the baseline hazard assumed, e.g. a constant, a piecewise constant, Weibull, Gompertz, or extreme value distribution.
Assuming the simplest case, that the hazard above is constant over time, i.e. not dependent on t, we get following hazard and survival function

$$\lambda_i(t|x_i) = \lambda_0 exp(\beta x_i)$$

$$S_i(t|x_i) = exp\big(-\lambda_0 t exp(\beta x_i)\big) = exp(-\lambda_0 t)^{exp(\beta x_i)}$$

Replacing λ_0 with $exp(\beta_0)$ in the hazard function above we get an expression that reminds of a transformed linear model, called exponential regression model.

$$\lambda_i(t|x_i) = exp(\beta_0 + \beta x_i)$$

GENERAL HAZARD RATE MODEL

We may want to allow predictors X to be time dependent. Examples of time-dependent predictors could be e.g., blood samples of a specific biomarker taken over time, where we are interested in studying the most recent value, the average of available values last three months, or variability of values last year, as the biomarker's impact on outcome.

We may also want to allow β coefficients to be time-dependent, implying that proportional hazard is no longer assumed. This enables the possibility to study changes in effects on an outcome over time for different variables. E.g., a certain treatment effect may be large closer to the initiation of the medication, but this effect might decrease after a certain time while the underlying disease is progressing. Alternatively, a treatment might need some time to accumulate in the body to affect an outcome. In both examples the treatment effect varies with time.

Such a model may be seen as a general hazard rate model and expressed by

$$\lambda_i\big(t|x_i(t)\big) = \lambda_0(t) exp\big(\beta(t) x_i(t)\big)$$

Note that both β coefficients (time-dependent effects) and x_i (time-dependent predictors) may vary with t.

There are three ways of fitting a general hazard rate model:

- Using a parametric approach by assuming e.g., exponential distribution for the $F(t)$ mentioned earlier.
- Using a semi-parametric approach, such as dividing time into small intervals and assuming that the baseline hazard is constant in each interval. This model is called piecewise exponential model, described in more detail below.

- Using a non-parametric approach that leaves baseline hazard completely unspecified, corresponding to the Cox model previously described. This model is also equivalent to the piecewise exponential model where time is modelled *ad absurdum* with one parameter per unique event time.

PIECEWISE EXPONENTIAL AND EXTENDED POISSON MODEL

In this method we partition the time t into J intervals, $0 = \tau_0 < \tau_1 < \ldots < \tau_J = \infty$. The j-th interval is defined as $[\tau_{j-1}, \tau_j]$. Assuming to have baseline hazard constant within each interval, i.e., not dependent on time, we get

$$\lambda_0(t) = \lambda_j \text{ for each } [\tau_{j-1}, \tau_j]$$

This implies that the baseline hazard is modelled using $\lambda_1, \lambda_2, \ldots, \lambda_J$, called piecewise constant model or piecewise exponential model.

Introducing predictors and proportional hazards assumption, we get

$$\lambda_{ij} = \lambda_j exp(\beta x_i)$$

Taking logs on both sides, a log-linear additive model is obtained, where $\alpha_j = log\lambda_j$

$$log\lambda_{ij} = \alpha_j exp(\beta x_i)$$

The piecewise exponential model above was proven in 1980-1981 independently by Holford, and Laird and Olivier, to be equivalent to a certain *Extended Poisson model*, through its log-likelihood.[128-131] Applications using generalized linear models were later published by others.[132-134]

For each participant partitioned data rows are created. For each interval free of event the time $t_{ij} = \tau_j - \tau_{j-1}$ and for the interval where the event occurs at t_i the $t_{ij} = t_i - \tau_{j-1}$. Beside interval-specific time points the indicator of the event, d_{ij}, is also created for each interval, being 0 for all intervals except the one in which the event occurs, where it is set to 1. Treating the event indicators d_{ij} as independent Poisson observations a piecewise exponential model is fitted for their means

$$\mu_{ij} = t_{ij}\lambda_{ij}$$

Taking the log on both sides, and expressing $\alpha_j = log\lambda_j$, we get

$$log\mu_{ij} = logt_{ij} + \alpha_j + \beta x_i$$

meaning that this model is equivalent to a Poisson log-linear model for partitioned observations with d_{ij} as the outcome and $logt_{ij}$ as an offset.

The proof of the two models being equivalent is because the likelihoods for the two coincide and therefore solve the same statistical problem and provide the same β estimates.

Both time-dependent predictors and time-dependent effects may be incorporated similarly as previously shown, if the interval width is large enough to account for the variable and effect changes. If not, smaller partitioning might be performed if required to update a variable over time. The limitation of this method is that it can be time-consuming for large datasets, particularly time-dependent models.

This general hazard rate model, allowing study of time-dependent predictors and time-dependent effects, *Piecewise exponential model*, or *Extended Poisson model*, was used to develop DIGIROP-Birth prediction models in *Paper I* and *Paper IV*.

2

2 RATIONALE

Prediction of risk is frequently studied in medical research. The development and validation of ROP prediction models have been ongoing for the past 20 years. Following the review of *American Academy of Ophthalmology* in 2016 of 23 studies, Hutchinson et al. concluded that more rigorous efforts must be implemented in prognostic ROP research.[135] Most often, the prediction models are based on the linear relationship between predictors and the outcome. Non-linear associations, interactions and time-updated information are seldom considered.

Advances in medicine and healthcare have, over years, implied an increase in survival of preterm infants with decreasing gestational ages. Infants born at GA less than 24 weeks are part of today's national patient registry data, surviving due to improved intensive medical care.[48] Consequently, an increased number of infants need ROP screening examinations. Studies have shown that the infant's discomfort caused by the physical manipulations of the eye globe and procedures during examinations leads to significant distress of the infant, including changes in blood pressure and oxygen saturation that may impact on their cognitive development.[136-138] Hence, efforts must be made to learn more about the causes of stress for this population and we must find new methods to decrease harm and pain during examinations. An effort to decrease the number of unnecessary examinations without compromising infant safety, while maximizing utilization of the healthcare resources is also critical. The clinical goal of a validated prediction tool within this project is to identify infants that do not need screening as well as to identify the time point when the ophthalmologist can safely stop screening examinations and hence reduce the number of unnecessary ROP examinations in this fragile population.

3

3 AIM

This thesis aimed to develop and validate prediction models for severe ROP requiring treatment using easy-obtainable variables available for all children nationally without compromising their outcome. Additionally, it was to describe the hazard function for ROP treatment and demonstrate the prognostic value of days on parenteral nutrition on any ROP and ROP requiring treatment. The outcome from the models was to be compared to other most cited ROP prediction models. Specifically, the aims were:

PAPER I **DIGIROP-Birth 1.0:** To develop, and internally and externally validate the DIGIROP-Birth model based solely on birth characteristics, including ~7300 infants from SWEDROP (2007-2018), ~1500 infants from the US (2005-2010) and ~300 infants from Germany (2011-2017). Further, it was to describe the hazard function for ROP treatment for infants born at <31 weeks of gestation. The models were to be compared to WINROP, CHOP-ROP, OMA-ROP and CO-ROP.

PAPER II **DIGIROP-Screen 1.0:** To develop and internally and externally validate DIGIROP-Screen models using DIGIROP-Birth risk estimates and data from ROP screening for infants born at 24-30 weeks of gestation, based on ~7300 infants from SWEDROP (2007-2018), ~600 infants from the US (2006-2009 and 2014-2019) and 300 infants from Germany (2011-2017). Additionally, a clinical decision support tool that achieved 100% sensitivity and the highest specificity possible was proposed. The models were to be compared to WINROP, CHOP-ROP, OMA-ROP, and CO-ROP.

PAPER III **External validation:** To externally validate DIGIROP-Birth, DIGIROP-Screen, and their clinical decision support tool on a contemporary Swedish cohort of ~1000 infants from SWEDROP (2018-2020).

PAPER IV **DIGIROP-Birth 2.0 and DIGIROP-Screen 2.0:** To demonstrate the prognostic value of days on parenteral nutrition on any ROP and ROP treatment. Additionally, it was to update and validate DIGIROP prediction models and their clinical decision support tool, including ~11,000 ROP-screened infants irrespective of GA from SWEDROP (2007-2020) incorporating days on parenteral nutrition, and BW instead of BWSDS. The models were to be compared to WINROP and G-ROP.

4

4 ETHICAL APPROVALS

Ethical permits have been obtained for all studies included in this thesis. All four studies are retrospective, register studies, including data retrieved from the medical records, considering routine examinations. At the first ROP screening examination, the parents are given the possibility to decline the infant's participation in the registry, following information provided orally and in writing about SWEDROP.

PAPER I The ethics committee approved the study at Uppsala University (original application Dnr 2010-117, and amendment Dnr 2010-117/2). Concerning the international data for validation of the models, ethical approvals were available for German and US data from the respective ethical committee and institutional review boards from all participating centers.

PAPER II The ethics committee approved the study at Uppsala University, Uppsala, Sweden (original application Dnr 2010-117, and amendment Dnr 2010-117/2). Data validation required a review of medical records approved by the Swedish Ethical Review Authority (Dnr 2019-02321). Concerning the international data for validation of the models, ethical approvals were available for German and US data from the respective ethical committee and institutional review boards from all participating centers.

PAPER III The Swedish Ethical Review Authority approved this study (Dnr 2019-02321). Ethical approval was available for extraction of data from SWEDROP until 31 December 2019 (original application Dnr 2010-117, and amendment Dnr 2010-117/2), and separately for the two regions during the year 2020 in the amendment (Dnr 2020-06940) to the original study application (Dnr 2019-02321).

PAPER IV The Swedish Ethical Review Authority approved this study (Dnr 2019-02321, an amendment to extend the study for years 2007-2025 Dnr 2022-02656-02). Ethical approval was available for data extraction from the SWEDROP until 31 December 2025 (Dnr 2021-05134) based on approvals Dnr 2010-117 and Dnr 2010-117/2.

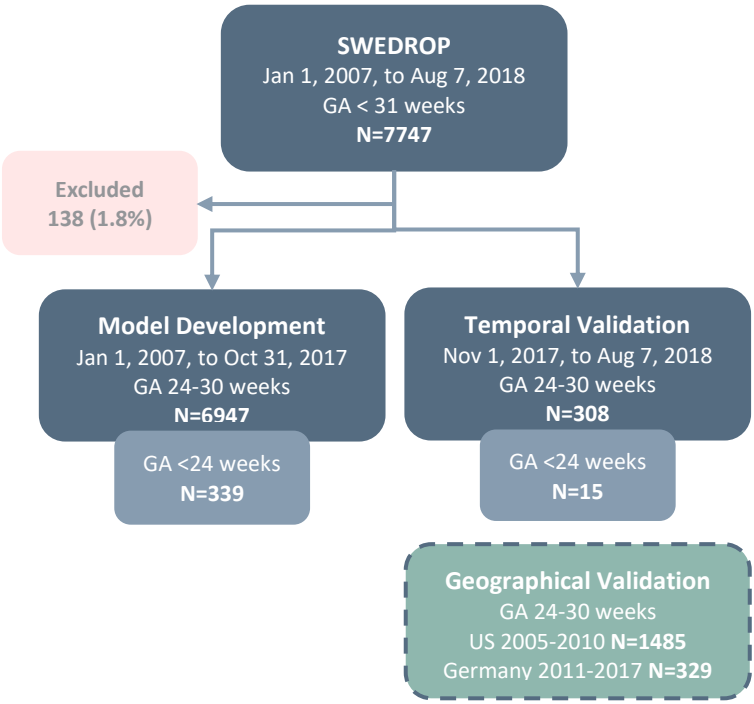
5

5 MATERIALS AND METHODS

5.1 PAPER I DIGIROP-BIRTH

The development of the DIGIROP-Birth 1.0 prediction model was based on the infants registered in SWEDROP 2007-2017. External validation included infants from SWEDROP 2017-2018, the US from 2005-2010, and Germany from 2011-2017.

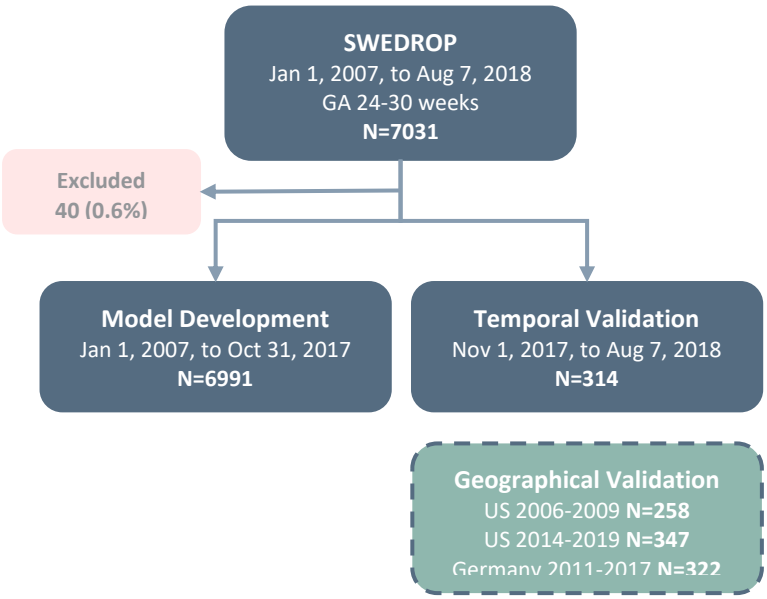
FIGURE 6. THE STUDY POPULATION INCLUDED IN *PAPER I*.



5.2 PAPER II DIGIROP-SCREEN

Development of the DIGIROP-Screen 1.0 prediction model was based on the same data extraction from SWEDROP 2007-2017 that was used for DIGIROP-Birth 1.0. In between the two studies, missing and incomplete data was validated based on the medical records. External validation included data from SWEDROP 2017-2018, two cohorts from the US 2006-2009 and 2014-2019, and one from Germany 2011-2017.

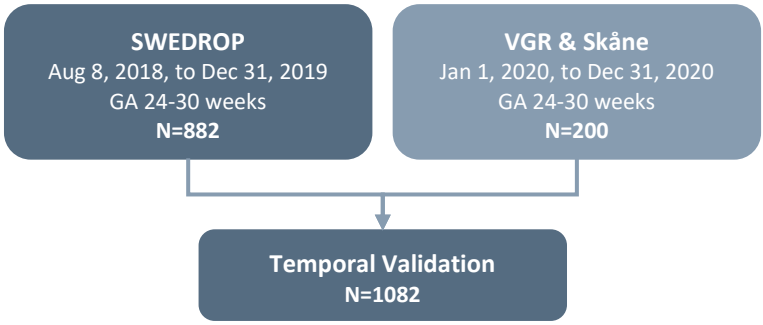
FIGURE 7. THE STUDY POPULATION INCLUDED IN *PAPER II*.



5.3 PAPER III EXTERNAL VALIDATION

External validation of DIGIROP-Birth 1.0, DIGIROP-Screen 1.0 and their clinical decision support tool was based on a contemporary extraction of data from SWEDROP 2018-2019 and regional data from Västra Götaland region (VGR) and Skåne 2020. There were no missing data.

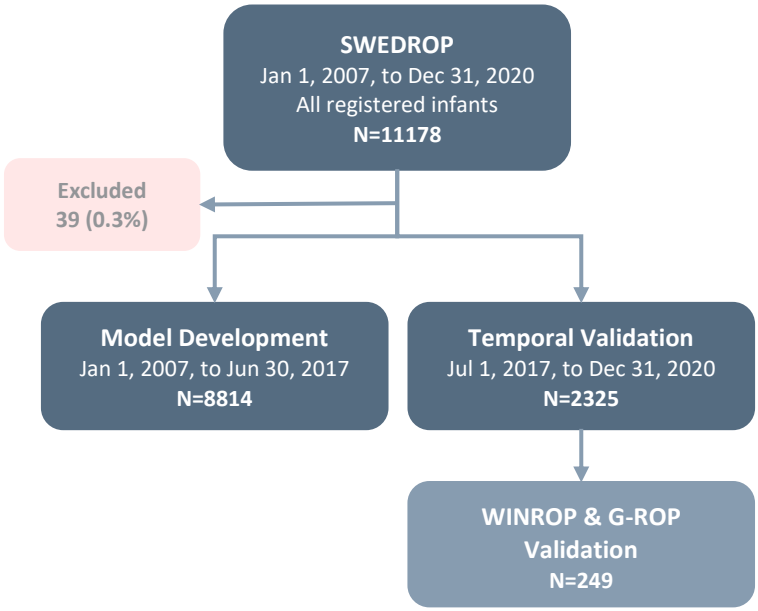
FIGURE 8. THE STUDY POPULATION INCLUDED IN *PAPER III*.



5.4 PAPER IV UPDATED DIGIROP MODELS

Update of DIGIROP-Birth 2.0, DIGIROP-Screen 2.0, and their clinical decision support tool was based on data extracted from SWEDROP between 2007 and 2020. This study included all registered infants without any restrictions on GA.

FIGURE 9. THE STUDY POPULATION INCLUDED IN *PAPER IV*.



5.5 OUTCOMES

The prediction models were developed to estimate the risk for severe ROP requiring treatment, defined either per the ETROP criteria, or based on the assessment of the examining ophthalmologist.[92]
In *Paper IV*, the association between PND and any ROP was assessed besides ROP treatment.

5.6 PREDICTORS

Predictors required for calculations of DIGIROP-Birth and DIGIROP-Screen risk estimates are presented in ***Table 3*** below.
Gestational age was ascertained by the fetal ultrasound that in Sweden is routinely performed in gestational week 18-20. PNA, PMA and GA were defined by applying the policy issued by the *AAP*.[1]
Parenteral nutrition duration considers total number of days on protein and lipid parenteral supplementation.

TABLE 3. PREDICTORS INCLUDED IN DIGIROP-BIRTH AND DIGIROP-SCREEN PREDICTION MODELS. *ROP = retinopathy of prematurity; SDS = standard deviation score.*

Variables	DIGIROP-Birth 1.0	DIGIROP-Birth 2.0	DIGIROP-Screen 1.0	DIGIROP-Screen 2.0
Gestational age (weeks and days)	X	X		
Birth weight (grams)		X		
Birth weight SDS[38] (z-score)	X			
Sex (boy, girl)	X	X		
Parenteral nutrition duration (<14 days, ≥14 days, unknown)		X		
Status about first ROP diagnosis (yes, no)			X	X
Age at first ROP diagnosis (weeks)			X	X
DIGIROP-Birth 1.0 risk estimate			X	
DIGIROP-Birth 2.0 risk estimate				X

Additionally, for DIGIROP-Screen 1.0, location (nasal, temporal, nasal and temporal) was investigated as a predictor. Given the uncertain reporting of this variable, and incomplete data, it was decided to leave this variable out of the final model. Time since the first detection of ROP was also evaluated and found redundant in the model when age at first ROP diagnosis was included.

5.7 GUIDELINES

The *Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis* (TRIPOD) statement was followed in all studies.[139] In *Paper I*, the *Prognosis Research Strategy* (PROGRESS) 3 was followed.[140] In *Paper III* and *Paper IV*, the *Prediction model study Risk of Bias Assessment Tool* (PROBAST) instrument was additionally applied.[141 142]

5.8 STATISTICS

SAMPLE SIZE

Data available in Sweden for *Paper I* included ~7000 infants born at GA 24 to 30 weeks. Among those, ~300 infants reached the endpoint, severe ROP requiring treatment. The number of 300 treated infants allows for a lower 95% CI of the required 100% sensitivity to be at least 99%, considered necessary by researchers in this field.[135 143] Of note is that most (21 out of 25) of the up-to-date published ROP prediction models are based on data samples of <2000 infants. The most extensive data set currently available worldwide is collected by a North American research group, including ~11,500 prematurely born infants.[144] Compared to other studies and available cohorts, it was evaluated that prediction models based on ~300 events meet the sample size requirements. In *Paper IV*, ~11,000 infants could be included from SWEDROP owing to more extended inclusion of calendar years and extension of the models to include all ROP-screened infants.

GENERAL METHODS

Database programming and statistical analyses were performed in SAS software version 9.4 (SAS Institute Inc., Cary, NC, USA). Tables and figures for the publications and the thesis were created in SAS software. Extended Poisson models for *Paper IV* were run in SAS software and validated in R Statistical software version 4.2.0 (R Core Team 2022, Vienna, Austria). The mathematical expressions were written in R Markdown.

Continuous variables were described by mean, standard deviation (SD), and/or median and range for descriptive purpose. Counts and percentages described categorical variables.

For comparisons between two groups Fisher's exact test was used for dichotomous variables, the Mantel-Haenszel Chi-square trend test for ordered categorical variables, the Chi-square test for non-ordered categorical variables, and Mann-Whitney U-test for continuous variables. The Jonckheere-Terpstra test was applied when studying the relationship between a continuous and an ordered categorical variable. The correlation between two continuous variables was described by Spearman correlation, r_S.

All tests were two-sided and a p-value <0.05 was considered statistically significant. Missing data were not imputed.

EXTENDED POISSON MODEL

METHODOLOGICAL HISTORY RELATED TO THIS THESIS

Development of DIGIROP-Birth models was performed applying extended Poisson model, i.e., piecewise exponential model, using an in-house built SAS macro transcribed and further developed by Artemis Mårtensson (published as Anton Mårtensson), MSc, at Statistiska konsultgruppen, Gothenburg, Sweden, during years 2010-2012. This macro is owned by Statistiska konsultgruppen and is not publicly available. Statistiska konsultgruppen gave permission to use the macro for this thesis. The transcription to SAS was based on a BASIC macro developed by Prof. Emeritus Anders Odén. The transcription work was performed in close cooperation with Prof. Odén and was initiated for a Safety-

GH project led by Sr. Prof. Kerstin Albertsson-Wikland.[145] Thorough validations were performed. Beside applications in other projects, Prof. Odén used the same method for the development of a statistical application known as FRAX within the field of osteoporosis, in which 10-year probability of fracture is calculated based on the hazard functions.[146]

During this thesis work, publications by Holford, Laird, and Olivier, on the proofs for the equivalence between piecewise exponential models and this extended Poisson model were identified.[128] [129] Whereupon the application of the method in SAS software was internally validated by its application in R software based on Prof. Bendix Carstensen's publication *Who needs the Cox model anyway?*.[147] Validations were performed for different research questions using several cohorts. The updated DIGIROP-Birth 2.0 model in *Paper IV* was developed using SAS macro and validated in R software using the *glm* function.

DEVELOPMENT

The development part of the DIGIROP-Birth model was divided into steps presented below.

STEP 1 DIGIROP-Birth 1.0	Model the hazard function for ROP treatment, first without including any predictors. For this purpose, a graphical representation of Epanechnikov's hazard was obtained. The hazard was modeled using various breakpoints in PNA. It was shown to increase around 8 weeks of PNA, peaked at around 12 weeks, and then decreased.
STEP 2 DIGIROP-Birth 1.0	A check of whether the same hazard shape was valid for GA, sex, and BWSDS categories was performed in a similar way as above. The hazard shape was concluded to be similar for GA, sex, and BWSDS categories.
STEP 3 DIGIROP-Birth 1.0	Various breakpoints for follow-up time were tested. Finally, a model including break-points at 8 and 12 weeks of PNA was chosen to define the shape of the underlying hazard.
STEP 4 DIGIROP-Birth 1.0	Similarly, using graphical presentation, a break-point at 27 weeks of GA and at -1 SDS for BW were selected for the model.

STEP *5* DIGIROP-Birth 2.0	In the updated DIGIROP-Birth 2.0 model BW was tested to be included instead of BWSDS. The Akaike's Information Criterion (AIC) showed only slightly better goodness-of-fit for BWSDS on infants born at GA 24-30 weeks. The decision was taken to use BW instead of BWSDS to include all infants registered in SWEDROP. BWSDS reference is not available for infants born at <24 weeks of gestation. In this updated model, PND categorized into <14 days, ≥14 days, and unknown (missing data, to not exclude any infants) was also included to account for infant morbidity status and negative impact of longer PND on ROP.
STEP *6* DIGIROP-Birth 1.0 DIGIROP-Birth 2.0	Contribution to the model by all possible interactions was tested, applying forward and backward selection. Those that had a p-value <0.05 were included in DIGROP-Birth 1.0. In DIGROP-Birth 2.0 the model with the lowest AIC was selected that corresponded to that following p<0.05.
STEP *7* DIGIROP-Birth 1.0 DIGIROP-Birth 2.0	Hazard functions, survival functions, and risk accumulated up to 20 weeks of PNA are calculated and expressed as model results.

In the analysis, the infants were followed until their first ROP treatment or until they were censored at 50 weeks of PMA. The method provided parameter estimates β, standard errors (SE) and p-values. HRs with 95% CI were estimated. A receiver operating characteristic (ROC) curve analysis using estimated risks for ROP treatment was performed and area under the curve (AUC), also called *c-statistics*, was presented. An AUC of 0.7-<0.8 is considered as acceptable, 0.8-<0.9 excellent, and ≥0.90 outstanding.[148]

In ***Appendix 1***, DIGIROP-Birth 1.0 and DIGIROP-Birth 2.0 models are presented, together with the calculations of the hazard function. Additionally, R code is provided for DIGIROP-Birth 2.0 model.

The survival function $S(t)$ was obtained using numerical integration, which was then used for calculation of its complement, cumulative distribution function $F(t)$. The 95% CI for $F(t)$ were obtained through 1000 repeated samples using a multivariate normal distribution of the parameter estimates and the covariance matrix from the extended Poisson model.

TECHNICAL DETAILS

To simplify interpretation of the interactions with GA, this variable was centered at 28 weeks, i.e., the variable used in the models was the translated *GA-28* instead. For the same reason, in the updated DIGIROP-Birth 2.0 *male* was coded as 0 and *female* as 1, while in DIGIROP-Birth 1.0 *male* was coded as 1 and *female* as 2. Estimates and HR for BW if DIGIROP-Birth 2.0 was expressed by 100 g increase. Re-naming and translating variables do not have any impact on the final results, i.e. the estimated probabilities. Such features might be introduced for the purpose of facilitating the interpretation of the parameter estimates during development.

In the published models, piecewise linear functions of time, GA and BWSDS were used, rather than splines, to easier express HR per week and per SDS increase, respectively, and for the interpretations of the interactions. Continuous variables are re-parametrized and divided into several variables depending on the number of break points applied, see ***Table 4*** below.

TABLE 4. RE-PARAMETRIZATION OF CONTINUOUS VARIABLES BEFORE BEING INCLUDED IN THE MODELS TO ACCOUNT FOR NON-LINEAR ASSOCIATIONS TO THE OUTCOME. *VAR = variable; SDS = standard deviation score; HR = hazard ratio.*

Variable (VAR)	Interval	Interpretation of HR	Re-parametrized into variables
Postnatal age (weeks)	[min, 8]	Per 1 week increase up to 8w	min(VAR,8)
	(8,12]	Per 1 week increase between 8 and 12w	min(max(VAR-8,0),12-8)
	(12,max]	Per 1 week increase after 12w	max(VAR-12,0)
Gestational age (weeks) (centered at 28 weeks)	[min,27]	Per 1 week increase up to 27w	min(VAR,-1)
	(27,max]	Per 1 week increase after 27w	max(VAR-(-1),0)
Birth weight SDS[38] (z-score)	[min,-1]	Per 1 SDS increase up to -1 SDS	min(VAR,-1)
	(-1,max]	Per 1 SDS increase above -1 SDS	max(VAR-(-1),0)

Modeling time using splines instead would provide smoother hazard functions. This application and its correspondence to the piecewise linear model are shown in the Results section.

BINARY LOGISTIC MODEL

Risk estimates from the DIGIROP-Screen were obtained by applying binary logistic regression in nine different models for PNA weeks 6-14. The choice of starting the screening model at PNA 6 weeks was due to the infants being screened at the earliest this time. About 70% of the infants in SWEDROP that needed ROP treatment were treated before PNA 14 weeks. Due to this and the fact that infants requiring such extended screening follow-ups are not a subject for an early discharge from the screening, the models were not further updated after this time. Each PNA week model excluded infants that have already received their first ROP treatment, e.g., in the PNA 8 weeks model infants that were treated before 8 weeks were excluded. In this way, the parameter estimates were re-estimated over time reflecting an interaction effect between the time and time-updated variables. Originally, efforts were made to develop this model by using extended Poisson regression including time-updated variables and effects. However, it was difficult to obtain a stable model producing reliable results, mostly owing to the limited number of events and different interacting effects for the problem. This is why binary logistic regression for several models was chosen instead.

DIGIROP-Screen 1.0 and DIGIROP-Screen 2.0 models are presented in the appendix of *Paper II* and *Paper IV*. Estimated probabilities were automatically extracted from the SAS software, calculated as following

$$P(event = 1) = \frac{1}{1 + exp(-LC)}$$

where $LC = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k$ for k different predictors in the model.

Logistic regression operates in the *logit* scale, which is why the DIGIROP-Birth risk estimates (probabilities) were transferred back into this scale before being added to the model.

$$logit\ p = ln \frac{p}{1 - p}$$

Unadjusted and adjusted binary logistic regression models were also used to evaluate the association between PND and ROP, any and treated. ROC curve analysis was performed to decide on the PND cut-off for DIGIROP-Birth 2.0 model. The cut-off at 14 weeks was found to have the highest AUC, which was 0.73.

Hosmer-Lemeshow test was performed for goodness-of-fit test of the models. From the logistic regression odds-ratios (OR), 95% CI and p-values were presented together with the AUC.

CLINICAL DECISION SUPPORT TOOL

The clinical decision support tool was based on the DIGIROP-Birth and DIGIROP-Screen risk estimates. In the DIGIROP-Screen 2.0 lower 95% CI for the risk was used to increase the safety of the models further. Based on the development cohort, GA-specific cut-offs were identified to achieve 100% sensitivity, which was then used to calculate specificity and validate both sensitivity and specificity on the external validation cohort. The usual practice is to select only one cut-off. However, from the graphical representation of the risk estimates for the ROP-treated and non-treated infants, a decrease in risks by increased GA week was observed, as well as visible discrimination of treated and non-treated infants per GA week. Due to the large dataset available, GA-specific cut-offs that favored the specificity obtained from the models could be chosen. The cut-offs were identified for GA <24 weeks handled together, 24+0 to 24+6 as 24 weeks, 25+0 to 25+6 as 25 weeks, 26+0 to 26+6 as 26 weeks, 27+0 to 27+6 as 27 weeks, 28+0 to 28+6 as 28 weeks, 29+0 to 30+6 handled together, and infants born at GA ≥31 weeks handled together. The maximum allowable cut-off was 0.05, i.e., 5% risk for ROP treatment.

For the clinical decision support tool 1.0, for infants born 24 to 30 weeks of GA in *Paper II*, three outlying infants were identified when defining the cut-offs for DIGIROP-Birth 1.0. These infants were treated despite not fulfilling the treatment criteria for Type 1 ROP. The decision was taken to exclude these infants from the development of the models and further evaluations. In *Paper IV*, where BW replaced BWSDS, and all infants were included irrespective of GA, this problem was not envisaged. Hence, no exclusions were made in *Paper IV* for

the updated DIGIROP-Birth 2.0, DIGIROP-Screen 2.0, and their clinical decision support tool.

VALIDATION

INTERNAL VALIDATION

DIGIROP models were internally validated using a 10-fold cross-validation, where the development cohort was randomly sub-divided into 10 equally large sub-cohorts. For each sub-cohort, the model was trained on each combination of nine sub-cohorts and the validation on the remaining sub-cohort, resulting in each observation being used only once for validation and nine times for training. Sensitivity and specificity were obtained from the validated data. For DIGIROP-Screen 1.0 and 2.0 cumulative specificity was the main specificity measure presented. It considers an infant discharged from the screening with a current PNA model if the infant was discharged from the screening with any previous PNA models or with DIGIROP-Birth, reflecting the intended clinical process.

Calibration plots show observed proportion and 95% CI of infants with ROP treatment on the y-axis and the mean estimated probabilities from the DIGIROP models on the x-axis. This was performed for the model development cohort, the validated cohort from the cross-validation, and the external validation cohort. Additionally, ROC curves for the models were obtained.

EXTERNAL VALIDATION

The models' and the clinical decision support tool's transportability in time and geographically were made by external validations on the temporarily different SWEDROP data, and other US and German data, respectively; see section 5.1-5.4 for description of external validation cohorts. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, and AUC were presented, as well as calibration plots and ROC curves.

COMPARISON TO OTHER ROP MODELS

DIGIROP models were compared to other ROP models/criteria. WINROP, CHOP-ROP, OMA-ROP, CO-ROP in *Paper I* and *Paper II*, and WINROP and G-ROP in *Paper IV*.[99 101 107 110 114 116 144] These models/criteria are based on GA, BW, and some functions of the longitudinal weights. Information about hydrocephalus was required for the G-ROP screening criteria. No hydrocephalus cases were observed in the validation cohort for *Paper IV*.
In *Paper I*, sensitivity for the comparison model was obtained based on each respective model's published cut-offs. Applying the same sensitivity, the cut-offs were set for DIGIROP-Birth 1.0. Then, the specificities were calculated.
Predictors and technical details are listed below per model/criteria.

WINROP GA, BW, weekly weight gain (weekly ≤36 weeks of PMA)	WINROP alarms are based on the Shiryaev-Roberts approach to detect significant deviation from the expected weight gain. Risk scores for *Paper I* and *Paper II* were already available in the WINROP validation data set used in the publication by Wu et al.[149] For the US validation cohort 2014-2019 and the SWEDROP validation cohort in *Paper IV*, the online application was applied.[100] The cut-off 2+3 (screen) *vs.* 0+1 (do not screen) for the alarms was used for comparison to DIGIROP. In *Paper IV*, observations with missing weekly weights (13.3%), crucial for calculating the alarms, were evaluated needing screening.
CHOP-ROP GA, BW, weight gain (latest and penultimate weekly weight)	The probability was calculated as 1/(1+exp(-Risk score)), where Risk score = (-1.50) + (4.24 if GA=23) + (3.49 if GA=24) + (3.60 if GA=25) + (2.33 if GA=26) + (2.48 if GA=27) + (-0.0037) × (BW) + (-0.0186) × (weight gain rate). Weight gain rate was calculated as (mean of daily weights of the preceding week – mean of daily weights of the penultimate week)/7. Probability cut-offs 0.0034 and 0.0140 from the original and validation study were used for comparison to DIGIROP.
OMA-ROP GA, BW, weight gain rate (latest ≤36 weeks of PMA)	Risk score = (weight - BW)/(date for latest weight – date of birth+1), where weight at week 36 was used or the latest before ROP treatment. A cut-off of 23 g/day was used for comparison to DIGIROP.

CO-ROP GA, BW, weight gain up to 4 weeks of PNA	CO-ROP alarm was set to 1 if GA<31 weeks, BW ≤1500 g and weight gain from birth to postnatal week 4 ≤650 g. If all three parameters were non-missing the alarm was equal to 0. Alarm cut-off 1 (screen) *vs.* 0 (do not screen) was used for comparison to DIGIROP.
G-ROP GA, BW, daily weight gain 10-19, 20-29, 30-39 days, hydrocephalus	The G-ROP screening criteria deems infants needing screening if any of the requirements are fulfilled: GA<28 weeks, BW <1051 g, weight gain between 10-19 days <120 g, weight gain between 20-29 days <180 g, weight gain between 30-39 days <170 g, or if the infant has hydrocephalus. An update of the criteria requires screening if weight gain <180 g for any of the three age intervals. To obtain weights for postnatal days 10, 19, 20, 29, 30, and 39, linear interpolation was applied, when values were not available precisely on these days. The observations with missing weights (9.2%) were evaluated as needing screening. Need for screening *vs.* no need for screening as per above according to the originally published and the updated criteria were used for comparison to DIGIROP. In the thesis, G-ROP was compared to DIGIROP also for cohorts from *Paper I* and *Paper II*, using linear interpolation to obtain missing weights at 10, 19, 20, 29, 30, and 39 days of PNA.

6

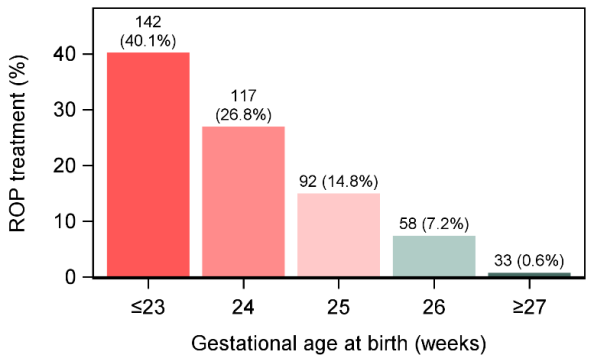
6 RESULTS

6.1 PAPER I DIGIROP-BIRTH

POPULATION

A total of 7609 infants from SWEDROP were included in *Paper I*; 55% were boys, the mean GA (SD) was 28.1 (2.1) weeks, and the mean (SD) BW was 1119 (353) g. Any ROP was developed in 32% and ROP treatment was provided to 442 infants (6%), 142 (40%) among those born at <24 weeks of GA, and 300 (4%) among those born 24-30 weeks of GA. In the external validation, 125 (8%) had ROP treatment in the US cohort, and 17 (5%) were treated for ROP in the German cohort. ***Figure 10*** below shows the GA-related incidence of ROP treatment among 7609 SWEDROP infants.

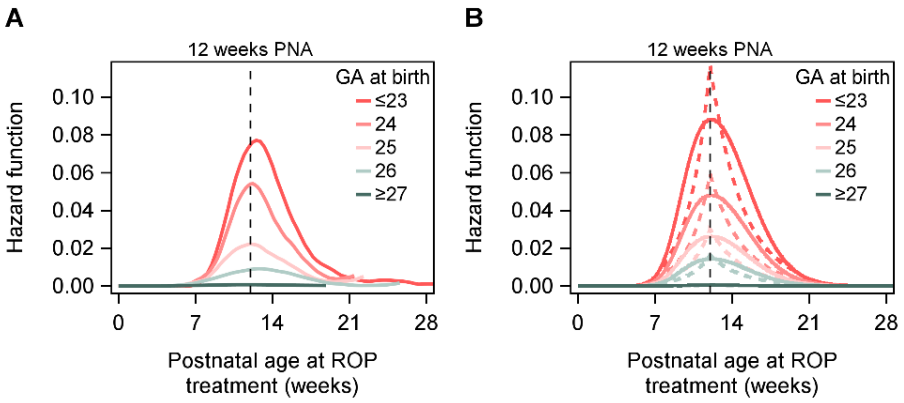
FIGURE 10. NUMBER AND PERCENTAGE OF INFANTS WITH ROP TREATMENT AMONG THOSE INCLUDED FROM SWEDROP 2007-2018 (N=7609). *ROP = retinopathy of prematurity.*



HAZARD FUNCTION FOR TREATED SEVERE ROP

In a simple extended Poisson model using time in study (PNA) adjusted for GA, the risk for ROP treatment increased by 54%, HR 1.54 (95% CI 1.39 to 1.70) per week from postnatal week 8 to 12, whereafter it decreased by 30%, HR 0.70 (95% CI 0.67 to 0.74). Below, the corresponding hazard function is presented using piecewise linear model and splines to illustrate correspondence.

FIGURE 11. HAZARD FUNCTION FOR ROP TREATMENT BY GESTATIONAL AGE ESTIMATED BY A) EPANECHNIKOV'S KERNEL SMOOTHING FOR GA STRATA, B) SPLINE (SOLID LINE) AND PIECEWISE LINEAR (SHORT-DASHED LINE) EXTENDED POISSON MODEL FOR MEAN GA WITHIN EACH GA STRATA. *GA = gestational age; PNA = postnatal age; ROP = retinopathy of prematurity.*



DIGIROP-BIRTH 1.0 FOR GA 24-30 WEEKS

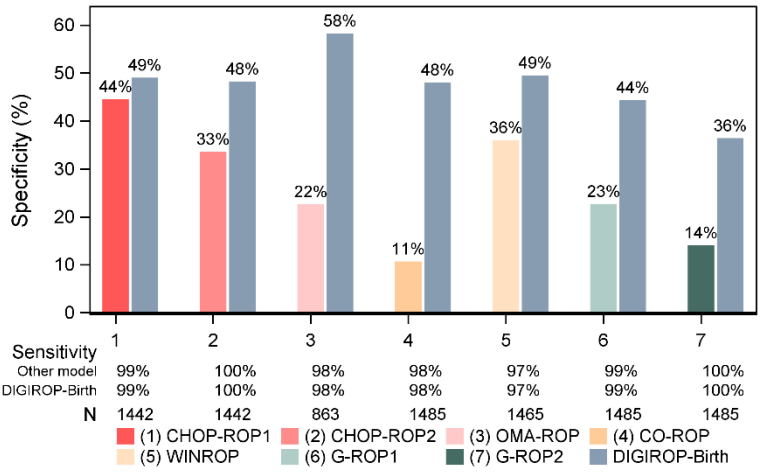
DIGIROP-Birth 1.0 requires data about GA, BW, and sex to estimate the risk for ROP treatment. It is developed for infants born at GA 24-30 weeks. As expected, lower GA and lower BWSDS showed a higher risk for ROP treatment. Interaction sex×GA was significant in the models indicating a greater decreasing risk for girls than for boys for increasing GA, HR at 25 weeks 0.83 (95% CI 0.64 to 1.07) and HR at 27 weeks 0.50 (95% CI 0.33 to 0.76), p for interaction 0.02. Other significant interactions were PNA×GA, and PNA×BWSDS. The internal and external validations showed AUC ranging between 0.87 and 0.94, with the observed *vs.* estimated probabilities for ROP treatment being well-distributed around the diagonal in the calibration plots.

The final model is presented in ***Appendix 1*** together with the hazard calculations.

DIGIROP-BIRTH 1.0 AND OTHER MODELS

Figure 12 below presents a comparison between DIGIROP-Birth 1.0 and selected ROP prediction models. Cut-offs for DIGIROP-Birth were not published in *Paper I*. Specificity for DIGIROP-Birth was based by identifying a cut-off for the same specificity as that obtained for other ROP models using their own published cut-offs. Therefore, for each ROP prediction model an own DIGIROP-Birth comparison was relevant. G-ROP criteria was added in the thesis. Weight at 10, 19, 20, 29, 30, and 39 days were obtained using linear interpolation of available data.

FIGURE 12. SENSITIVITY AND SPECIFICITY FOR DIGIROP-BIRTH 1.0 COMPARED TO OTHER ROP PREDICTION MODELS. *CHOP-ROP1 and G-ROP1 used the cut-offs published in their original publications, respectively, CHOP-ROP2 and G-ROP2 used the cut-offs published in their validation studies. Different numbers of infants were included in different comparisons.*



6.2 PAPER II DIGIROP-SCREEN

POPULATION

Of the 6991 infants from SWEDROP included in *Paper II* model development cohort, 55% were boys, the mean GA (SD) was 28.3 (1.9) weeks, and the mean BW (SD) was 1146 (339) g. Any ROP was developed in 29% in the development, and 41% in the validation cohort (n=1241). ROP treatment was performed in 287 infants (4%) in the model development cohort and in 49 (4%) infants in the validation cohort. The validation cohort had fewer boys, lower GA, lower BW, and more infants experienced any ROP than in the model development cohort.

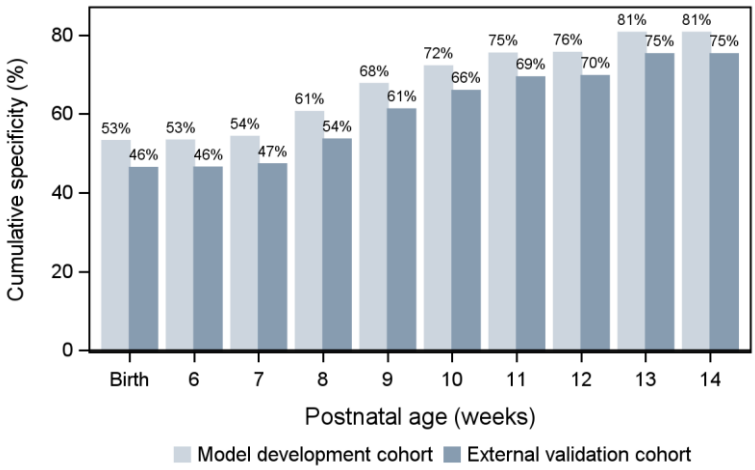
DIGIROP-SCREEN 1.0 FOR GA 24-30 WEEKS

DIGIROP-Screen 1.0 requires data about DIGIROP-Birth 1.0 risk estimate, status about first ROP diagnosis (yes/no) and PNA at first diagnosis of ROP to estimate the risk for ROP treatment. Nine different models over PNA 6 to 14 weeks allow for the models to update the parameter estimates used to estimate the risks for ROP treatment. By incorporating two-way and three-way interactions, models for infants with and without any diagnosis of ROP could be modeled together. For infants without an ROP diagnosis up to a certain PNA the intercept and DIGIROP-Birth estimates are the contributors for the risk. For those that do have an ROP diagnosis, age at the first diagnosis and interaction between this age and DIGIROP-Birth estimate are defining the DIGIROP-Screen risk estimate. The AUC for the models in the development cohort, internal cross-validation and external validation cohorts ranged between 0.88 and 0.94. Calibration plots showed well-calibrated models.

DIGIROP DECISION SUPPORT TOOL 1.0 FOR GA 24-30 WEEKS

For the required 100% sensitivity in the development cohort, the specificity for DIGIROP-Birth was 53%, and cumulative specificity over PNA weeks 6 to 14 increased from 53% to 81%. The respective figures for the external validation cohort were 46% at birth, and 46% to 75% during the screening. For all PNA models the sensitivity was 100% in the external validation except for one infant with a syndrome incorrectly deemed not needing screening at birth, and at PNA weeks 6 and 7.

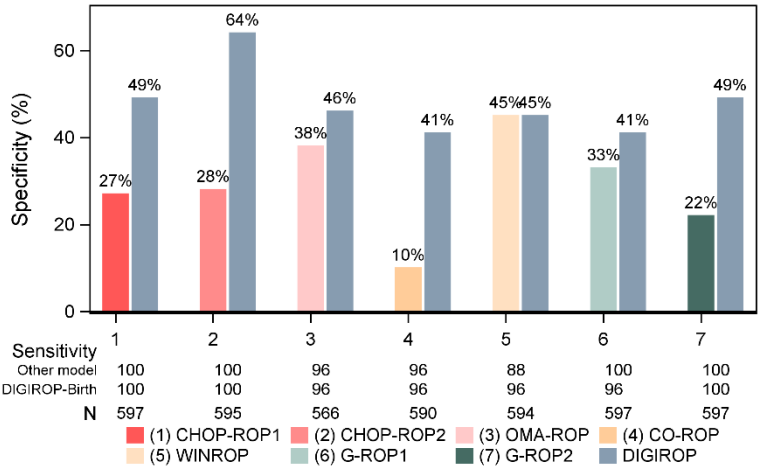
FIGURE 13. CUMULATIVE SPECIFICITY FOR DIGIROP DECISION SUPPORT TOOL 1.0.



DIGIROP-SCREEN 1.0 AND OTHER MODELS

Figure 14 below presents a comparison of DIGIROP-Birth and DIGIROP-Screen 1.0 with selected ROP prediction models.

FIGURE 14. SENSITIVITY AND SPECIFICITY FOR DIGIROP-BIRTH 1.0 AND DIGIROP-SCREEN 1.0 COMPARED TO OTHER ROP MODELS. *Comparisons: CHOP-ROP1 and DIGIROP-Screen PNA 8w, CHOP-ROP2 and DIGIROP-Screen PNA 12w, OMA-ROP and DIGIROP-Screen up to 36 weeks PMA, CO-ROP and DIGIROP-Birth, WINROP and DIGIROP-Screen up to risk flag or last measurement, G-ROP1 original criteria and DIGIROP-Birth, and G-ROP2 validated criteria and DIGIROP-Screen PNA 8w.*



6.3 PAPER III EXTERNAL VALIDATION

POPULATION

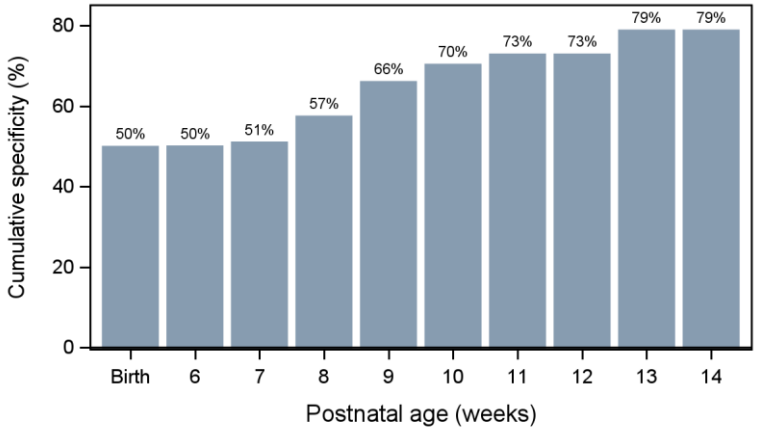
Of the 1082 infants from the contemporary extraction of SWEDROP infants (2018-2020) included in *Paper III*, 55% were boys, the mean GA (SD) was 28.2 (1.9) weeks, and the mean (SD) BW was 1117 (340) g. Any ROP was diagnosed in 31%. ROP treatment was performed in 57 infants (5%).

VALIDATION OF DIGIROP-BIRTH 1.0 and DIGIROP-SCREEN 1.0 FOR GA 24-30 WEEKS

The AUC for the DIGIROP models ranged between 0.93 and 0.97. The sensitivity for DIGIROP-Birth 1.0 was 96%, and the specificity was 50%. For DIGIROP-Screen 1.0 the sensitivity ranged between 93% and 100%, and cumulative specificity between 50% and 79%.

In total, 4 out of 57 infants were incorrectly classified as not needing screening. All four infants had severe comorbidities and/or had missed timely screening examinations. DIGIROP decision support tool was recommended not to be used for infants diagnosed with severe congenital malformations/syndromes, hydrocephalus, and those with performed intestinal surgery.

FIGURE 15. CUMULATIVE SPECIFICITY FOR DIGIROP-BIRTH 1.0 AND DIGIROP-SCREEN 1.0 IN THE CONTEMPORARY SWEDROP COHORT.



6.4 PAPER IV UPDATED DIGIROP MODELS

In *Paper IV*, the association between PND and ROP was studied. Further, DIGIROP models and the decision support tool were updated into 2.0 version, by including all ROP-screened infants registered in SWEDROP and adding PND as a predictor. Validations on external data and comparisons to other models were made. To avoid publication before manuscript acceptance, results for *Paper IV* are excluded from the thesis frame. All important information may be found in the *Paper IV* attached revised manuscript.

6.5 SIMPLIFIED MODELS FOR DIGIROP-BIRTH

Based on the data from *Paper IV* two simplified prediction models using linear relation of GA, BW, sex with and without days on parenteral nutrition (PND) excluding all interactions were developed and validated to visualize the differences *vs.* DIGIROP-Birth 2.0.

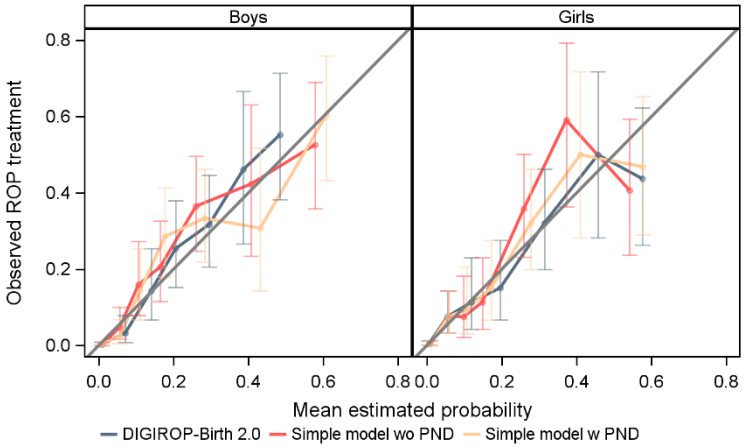
TABLE 5. SIMPLE PREDICTION USING GA, BW, SEX, WITH AND WITHOUT PND ON THE MODEL DEVELOPMENT COHORT. GA = *gestational age; BW = birth weight; PND = parenteral nutrition duration; HR = hazard ratio.*

Model/Variable	HR (95% CI)	p-value
Simple model without days on parenteral nutrition (PND)		
GA per 1 week increase	0.59 (0.54-0.64)	<0.0001
Sex (*girls vs. boys*)	0.80 (0.66-0.97)	0.022
BW per 100 g increase	0.79 (0.74-0.85)	<0.0001
Simple model with days on parenteral nutrition (PND)		
GA per 1 week increase	0.61 (0.56-0.67)	<0.0001
Sex (*girls vs. boys*)	0.81 (0.67-0.98)	0.031
BW per 100 g increase	0.82 (0.76-0.88)	<0.0001
PND (*≥14 days vs. <14 days*)	2.00 (1.56-2.58)	<0.0001
PND (*unknown vs. <14 days*)	1.28 (0.92-1.77)	0.14

Decision support tools using the same methodology as in the original models were developed for these two new models. For the simple model without PND, specificity in the model development cohort was 50% for the required 100% sensitivity. In the external validation cohort, a specificity of 41% was obtained, and 2 out of 152 infants with ROP treatment were incorrectly flagged not needing screening. The simple model including PND had a specificity of 49% in the model development cohort and 40% in the external validation cohort. All 152 infants with ROP treatment were correctly identified needing ROP screening. Similar results were obtained for the DIGIROP-Birth 2.0 including PND and significant interactions. This means that concerning only sensitivity and specificity the simple model including PND without interactions performs similarly to the more complicated model.

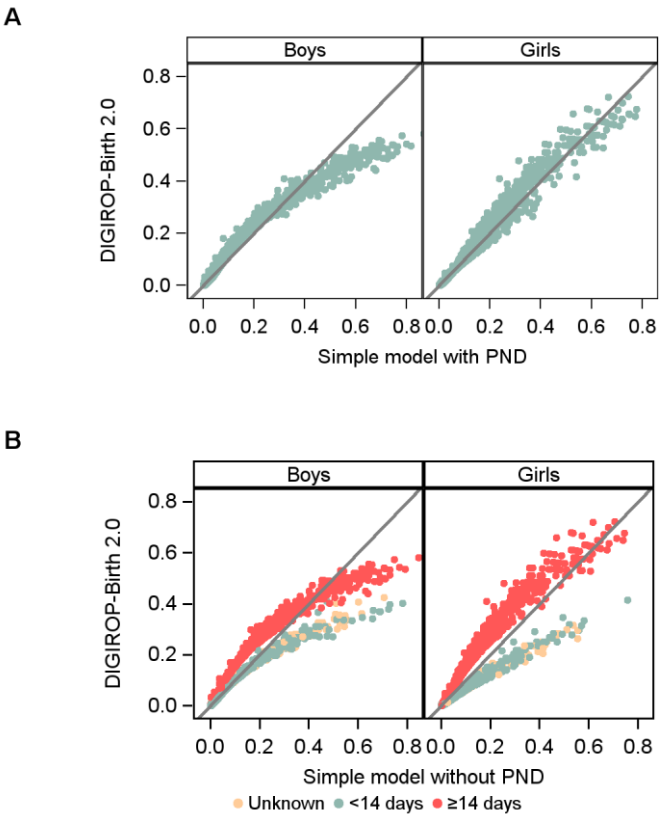
The calibration plots, showing how well probability estimates correspond to the observed ones in different regions, are given in ***Figure 16*** below. The simple model without PND performed less well for high-risk girls. For boys on the other hand the simple model with PND performed less well. In some regions the probabilities were underestimated and in others overestimated. Hence, including important predictors such as PND and interactions has been shown to improve the reliability and accuracy of the model.

FIGURE 16. CALIBRATION PLOTS FOR DIGIROP-BIRTH 2.0, SIMPLE MODEL WITHOUT PND, AND SIMPLE MODEL WITH PND BY SEX, PERFORMED ON EXTERNAL VALIDATION COHORT FROM SWEDROP 2017-2020. *PND = parenteral nutrition duration; ROP = retinopathy of prematurity; w = with; wo = without.*



In ***Figure 17.A*** the estimated probabilities for DIGIROP-Birth 2.0 and the simple model with PND are presented against each other. The underestimation for low-risk boys and overestimation for those with high risk for ROP treatment is observable, while the probabilities for girls are better distributed around the diagonal. ***Figure 17.B*** shows the corresponding probabilities for DIGIROP-Birth 2.0 *vs.* the simple model without PND. The importance of PND in the model is visible for both sexes but has greater impact on the probabilities for girls than for boys.

FIGURE 17. PROBABILITY ESTIMATES FOR A) DIGIROP-BIRTH 2.0 *VS.* SIMPLE MODEL WITH PND, B) DIGIROP-BIRTH 2.0 MODEL *VS.* SIMPLE MODEL WITHOUT PND, PERFORMED ON DATA FROM SWEDROP 2007-2020. *PND = parenteral nutrition duration.*



Given these results, we may conclude that both PND and the interactions (sex × PND shown above), are important to be included in the optimized model.

7

7 ONLINE APPLICATION

DIGIROP models are available free-of-charge in an online application at www.digirop.com.[150] The application was developed by the company Fooheads in cooperation with our research group, based on the provided risk estimations with 95% CI for DIGIROP-Birth, parameter estimates, covariance matrices and algorithms for DIGIROP-Screen, and figures based on the SWEDROP data.

FIGURE 18. DIGIROP-BIRTH 1.0 ONLINE APPLICATION.

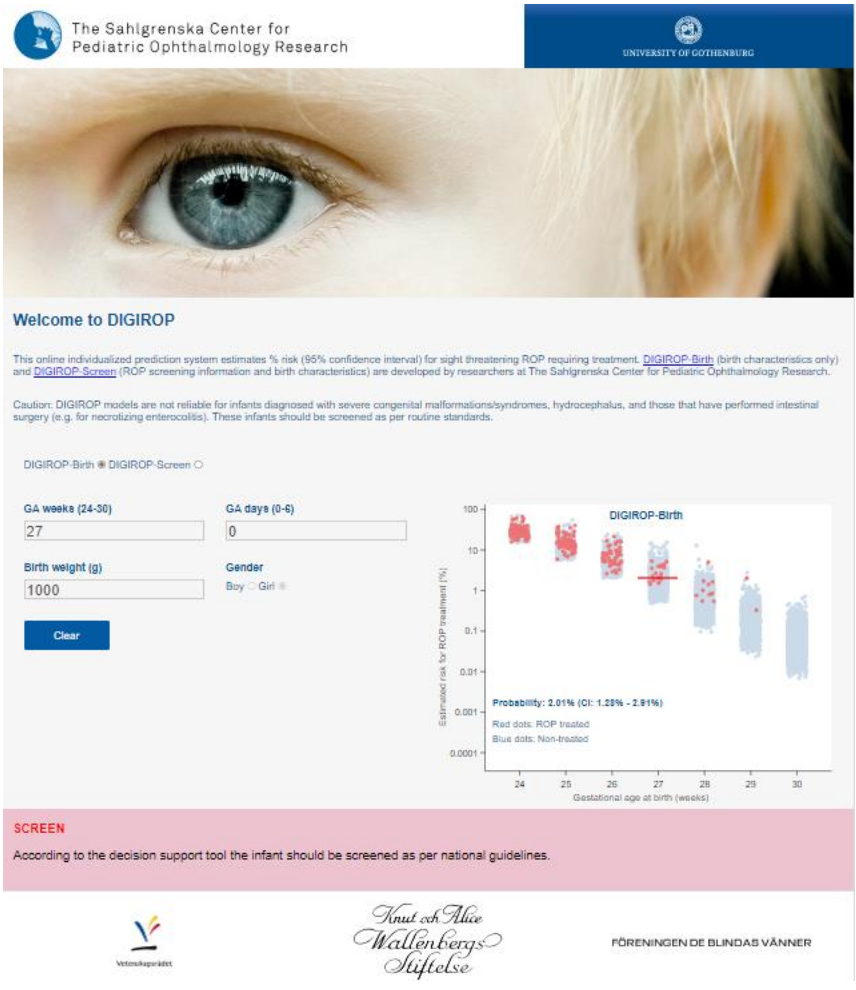


FIGURE 19. DIGIROP-SCREEN 1.0 ONLINE APPLICATION WITH ROP DIAGNOSIS BEFORE POSTNATAL WEEK 9.

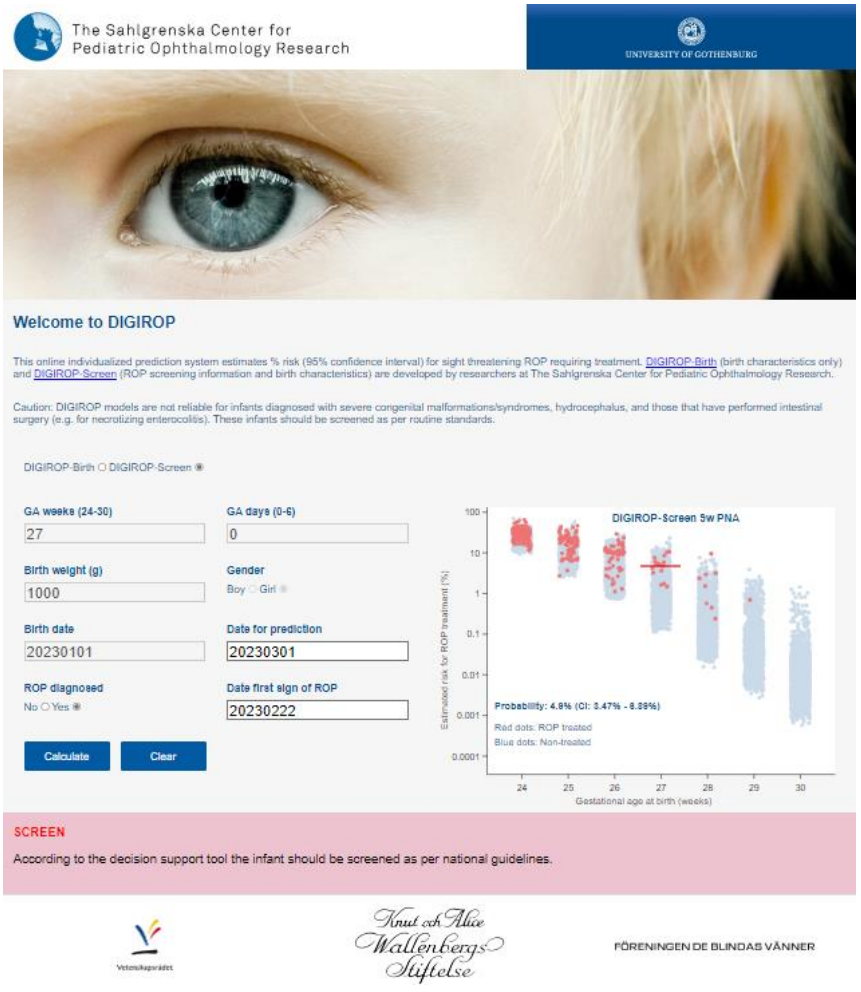
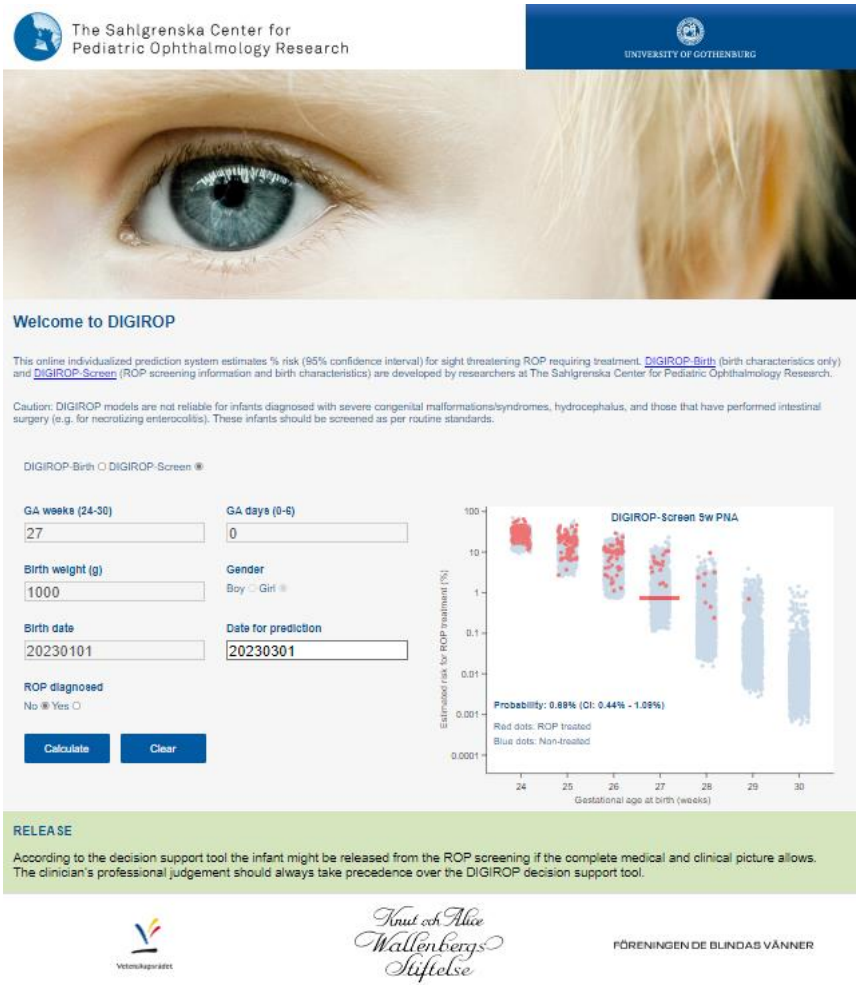


FIGURE 20. DIGIROP-SCREEN 1.0 ONLINE APPLICATION WITHOUT ROP DIAGNOSIS BEFORE POSTNATAL WEEK 9.



In ***Appendix 2***, information provided to researchers, requesting to use DIGIROP models, is presented.

8

8 DISCUSSION

DIGIROP models are developed to early identify low-risk infants that can be safely released for all or some unnecessary ROP screening examinations. Below, clinical implications, including the infant's well-being, economic benefits, selected statistical issues, selected results, and ethical considerations are discussed.

8.1 CLINICAL IMPLICATIONS

The estimated risk predictions from the DIGIROP models can be implemented in the clinic as one among other tools available to guide clinical recommendations and facilitate decisions. Nevertheless, it is critical that each infant's entire disease status is considered and that actions are taken based on the physician's complete medical assessment.

INFANT WELL-BEING

ROP screening with ophthalmoscopy is technically challenging and must be performed by a specialized and experienced ophthalmologist. Screening using objective imaging also requires a skilled person to take the images and an experienced ophthalmologist to interpret them. Before the examination, the infant's pupils are routinely dilated with mydriatic eye drops, and if images are taken, anesthetic drops are also given. Studies have shown that mydriatic drugs applied as eye drop have systemic effects, including the gastrointestinal side effects, apnea, bradycardia, and oxygen desaturation.[137 151 152] In addition, physical manipulation of the eye during ROP screening examinations is painful in neonates, and is associated with oxygen desaturation, increased heart rate, and infant distress.[136 137 153 154] Exposures to painful and stressful neonatal procedures may affect an infant's brain development.[138] Therefore, limiting the duration of examinations and using prediction models to reduce the number of unnecessary ROP examinations may contribute to improved well-being of infants.

PHYSICIAN WORKFLOW

The ophthalmologist's decision regarding the scheduling of ROP screening examinations is based on the recommendations given in the national guidelines for ROP screening.[65] There are various barriers to changing physician workflow, as were described in a study of the implementation of a prediction model in the clinic.[155] The important facilitators for the successful implementation of a prediction tool are automatic calculations of the risks, a directive approach to recommendations, prediction of relevant outcomes, and smooth integration of the prediction models in the existing clinical workflow. DIGIROP models are freely available, provide automatic calculations and propose decisions based on estimated risks, hence directing the ophthalmologist rather than assisting. The studied outcome considers the severity of ROP and is highly relevant to screening ophthalmologists. In implementing a prediction tool, it is preferable to minimize the number of additional working steps, and to integrate the models into the currently used systems for entering medical data. This action might be more or less difficult depending on the data safety and monitoring of the systems used for each health care provider. To assure the long-term safety of a preterm infant who is released from ROP examinations by the screening tool, clear measures for future communication between parents/guardians and the neonatology department with regard to the follow-up of the infant's development and changes in their well-being should be established. An application installation may facilitate this communication.

PARENT/GUARDIAN PERCEPTION

Not all individuals are expected to react in the same way when they find themselves in a crisis situation, such as the birth of a child born preterm. Suffering from prolonged stress can impair parent/guardian mental health[156] In addition, parent/guardian well-being and home environment have a significant impact on infant development.[157] Clear and consistent communication between healthcare providers and parents/guardians plays an important role in their management because of the increased stress levels.[158] Therefore, ophthalmologists must be sensitive to the wishes of parents/guardians and respect their decisions even when it concerns the receipt of the risk estimates obtained from the models. Their perception of the information provided by the models and the usefulness of such a tool may vary.

8.2 HEALTH ECONOMY

DIGIROP-Birth 2.0 achieved a specificity of 47% in the complete cohort of >11,000 infants for the required 100% sensitivity in the model development cohort. This corresponds to 44% of infants being released from all of the ~16,000 ROP examinations among those screened in the cohort in Sweden between the years 2007 and 2020. A total of ~10,000 ROP examinations were performed among 5% of infants that required ROP treatment, and ~55,000 examinations among the 95% that did not progress to treatment-requiring ROP. Approximately one-third of these 55,000 examinations could have been eliminated safely by applying DIGIROP models, which corresponds to the saved costs of up to 4 million US$ based on the estimated costs per one ROP screening examination in high-income economy countries. This is summarized in a review by Gyllensten et al.[159] Additionally, DIGIROP-Screen releases infants during the screening process that further increases this figure. However, the exact benefit in terms of health economics may only be obtained after the implementation of the tool in the clinics and evaluation of cost-effectiveness in two comparative groups with and without application of the clinical decision support tool.

8.3 STATISTICAL CONSIDERATIONS

Developing prediction models for a rare disease occurring in a rare population, requiring 100% sensitivity to avoid devastating outcomes, with the goal of optimizing specificity, is accompanied by some statistical challenges and unique features as discussed below.

THEORETICAL *VS.* REAL-LIFE SENSITIVITY AND SPECIFICITY

Validated on Swedish data, DIGIROP 1.0 models incorrectly classified four infants in *Paper III*. In practice, these infants would not have been released from screening as the clinical indication for ROP screening based on their severe medical conditions was fulfilled. DIGIROP 2.0 which added duration of parenteral nutrition, did correctly identify all high-risk infants requiring ROP treatment. Provided that no morbidity status, other than the proxy-variable prolonged duration of parenteral nutrition, is included in the DIGIROP risk estimations, physicians must be cautious regarding infants' medical status when

implementing this prediction tool in the clinics. In *Paper III*, we recommended that for infants diagnosed with severe congenital malformations/syndromes, hydrocephalus, and for those that have had intestinal surgery such as for NEC, routine ROP screening should apply. Therefore, in populations with similar levels of neonatal health care as is found in Sweden, DIGIROP 2.0 theoretical and real-life sensitivity are expected to reach 100%.

Statistically obtained specificity of 47% in the complete ROP-screened Swedish cohort during the years 2007-2020 presented in *Paper IV* may be considered as the theoretical upper limit for the specificity. In real life, following the implementation of the tool in the clinics, the specificity is expected to be lower. There are at least two reasons for that. The first is that the recommendations to apply routine screening for all infants meeting screening criteria based on their medical conditions are not considered in the models. The second reason is that clinicians are likely to initially have reservations concerning the use of the tool's clinical decision making, although these reservations will decrease over time with gained experience.

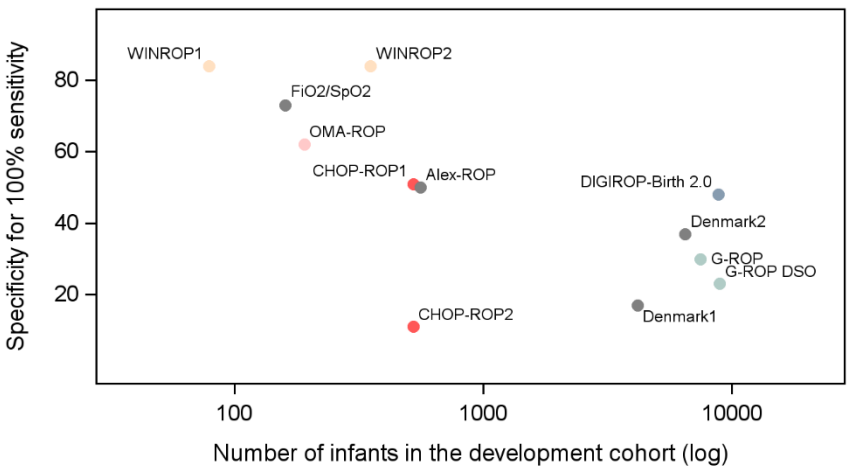
REQUIREMENT OF 100% SENSITIVITY

Although demanding 100% sensitivity is uncommon in prediction modelling research, it is necessary in the models developed for severe ROP needing treatment. Additionally, as recommended for ROP prediction models, the lower limit for the sensitivity should not be lower than 99%, corresponding to ~300 events, and ~6000 infants included in the development and validation considering a prevalence of 5%.[135 143] Such demands are difficult to meet, especially since the population of interest is a rare population. To date, only three ROP prediction models worldwide have included this number of infants in their development. These are the Denmark study, the G-ROP screening criteria, and DIGIROP models discussed here.[101 111 160-162]

We are strongly convinced that DIGIROP models and any models requiring 100% sensitivity will not always meet this demand in future validation studies, as opposed to real-life situations where infants' medical conditions are considered besides the tool's recommendation for clinical decision. The explanation is rather simple. The defined cut-offs are based on one certain infant or a few infants in the case of category-specific cut-offs. Given the fact that our outcome of interest is a

rare outcome, large validation cohorts are required in order to include an appropriate number of studied events, preferably ~100 events as recommended by the PROBAST instrument.[141] Validation studies performed on larger cohorts are disposed to more extreme or outlying infants than those performed on smaller cohorts, be it for the reason of specific cases of the disease, substandard medical judgments, or poorly entered data. Hence, although DIGIROP models are based on a complete Swedish population screened for ROP in the last 14 years, we cannot reject the possibility that even more extreme data will be observed sometime in the future. Additionally, larger development studies risk, due to the same reasons for including infants with outlying data, to have lower specificity if a high sensitivity of 100% or close to 100% is required. ***Figure 21*** below presents the relation between specificity and N (logarithm, to easier perceive the trend for N<1000) for ROP models with 100% sensitivity presented in ***Table 1***. The negative relation between the two variables is clearly visible.

FIGURE 21. RELATION BETWEEN SPECIFICITY AND NUMBER OF INFANTS INCLUDED IN THE DEVELOPMENT COHORT FOR REQUIRED HIGH SENSITIVITY.



One of the predictive ability measures for a model is the AUC, i.e., area under the ROC curve, that represents the predictive ability of the model over the complete ROC curve. DIGIROP models achieved AUC of >0.90, interpreted as an outstanding performance. For models requiring 100% sensitivity, achieving even higher AUC is of less importance. If the optimization of the specificity is of interest, as it is in our case, the most important feature of the curve is how fast it

increases up to 100% sensitivity, and this is dependent on the model's performance on certain outlying infants.

Positive and negative predictive value, PPV and NPV, are often considered more useful measures than sensitivity and specificity in daily clinical work. These measures are, however, dependent on the prevalence of the studied outcome as opposed to sensitivity and specificity. Models with rare events required to reach high sensitivity have very high NPV and very low PPV. If the model classifies an infant to be released from the screening, with the requirement of 100% sensitivity, we would be assured that NPV, the probability for this infant to truly not progress to severe ROP requiring treatment, would be 100% too. If the model, on the other hand, classified the infant to be screened, the PPV, expressing the probability for this infant to truly progress to severe ROP requiring treatment, would, in the case of rare diseases with 100% required sensitivity, be low. In our cohort of >11,000, an infant identified to be at high risk by DIGIROP-Birth 2.0 would have a ~10% risk of truly being a high-risk infant requiring treatment. The PPV of 10% applying the model may be compared to the PPV of 5%, that is, the prevalence, by not applying the model. In a clinical situation where the interventional act, here the ROP screening examination, is not directly hazardous for all infants undergoing these examinations and the price of blindness is much higher than the potential harm these examinations introduce, the low PPV is not a problem.

NON-LINEAR AND INTERACTING ASSOCIATIONS

The decision to include or not to include non-linear and interacting associations in a prediction model is dependent on the sample size available for the model development. Prediction models developed on small data sets requiring many parameters to be estimated imply *optimism* and an *over-parametrized* model, that, in external validation, risks performing much less well (all models are expected to perform less well in the external validation) than in the development cohort; *bias* risks to be introduced. In contrast, developing a too simple model on a large data set risks large *variance* for the estimates when important predictive variables are excluded. This so-called *bias-variance* trade-off needs to be taken into account during the development part. The additionally developed simple DIGIROP models, excluding non-linear and interacting associations presented in Section 6.5, showed that the model's reliability through calibration performance was

sacrificed compared to the more complex final model. According to the PROBAST instrument, the recommended number of events per variable (EPV) should exceed 20 to avoid introducing bias through over-parametrization.[141] For DIGIROP-Birth 1.0, EPV was 20.6 (289/14), and for DIGIROP-Birth 2.0, it was 31.9 (447/14).

8.4 SELECTED RESULTS

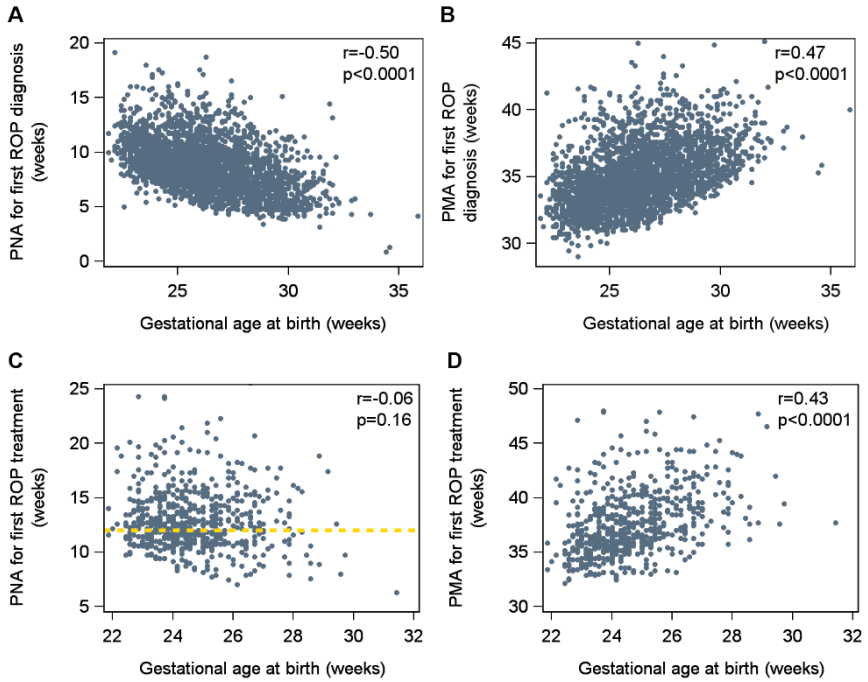
TIME COURSE OF THE DISEASE

In *Paper I* the time course of progression to severe ROP needing treatment, defining the instantaneous risk of this advanced stage of the disease, was described for the first time per GA at birth. The risk peaked at around 12 weeks' PNA irrespective of GA. This was confirmed by Holmström et al. 2019 using the same SWEDROP cohort for their publication that led to modification of the screening criteria in Sweden, showing the significant relation between GA at birth and PMA at ROP treatment, but not with PNA at ROP treatment.[48] Lack of correlation in this analysis means that the infant's age at ROP treatment is not differing for different GAs. Previously, landmark studies, and the clinical work, were not giving much attention to PNA, only PMA. The ETROP study described that the PMA was associated with the progression of pre-threshold ROP, as did the Cryotherapy for ROP (CRYO-ROP) study before that, meaning that the incidence of pre-threshold ROP differs with PMA.[163] [164] However, these studies have not evaluated the relationship with PNA, or the instantaneous risk of the disease.

In 1992, Quinn et. al., published a paper studying onset of ROP in relation to PNA and PMA.[165] This study showed that both PNA and PMA at the onset of ROP were associated with GA, indicating that both PNA and PMA are important to consider for the initiation and scheduling of ROP screening examinations. Using the cohort from *Paper IV*, we could replicate these results, as shown in ***Figure 22.A*** and ***Figure 22.B***. Infants with lower GA at birth develop ROP at earlier PMA but at later PNA than those born with higher GA. However, what was not studied by Quinn et al., was PNA and PMA at ROP treatment in relation to GA. Interestingly, as shown in *Paper I*, PNA at ROP treatment did not correlate with GA, but PMA at ROP treatment did. The results were confirmed using the

complete cohort from 2007-2020 including all ROP-screened infants reported in SWEDROP, ***Figure 22.C*** and ***Figure 22.D***. This suggests that the severe stages of ROP are highly dependent on the event of premature birth itself. Since this observation is independent of GA, a type of programming of disease progression from birth is insinuated.

FIGURE 22. RELATION BETWEEN A) POSTNATAL AGE AT THE ONSET OF RETINOPATHY OF PREMATURITY AND GESTATIONAL AGE, B) POSTMENSTRUAL AGE AT THE ONSET OF RETINOPATHY OF PREMATURITY AND GESTATIONAL AGE, C) POSTNATAL AGE AT THE FIRST RETINOPATHY OF PREMATURITY TREATMENT AND GESTATIONAL AGE, B) POSTMENSTRUAL AGE AT THE FIRST RETINOPATHY OF PREMATURITY TREATMENT AND GESTATIONAL AGE. *PNA = postnatal age, PMA = postmenstrual age, ROP = retinopathy of prematurity.*



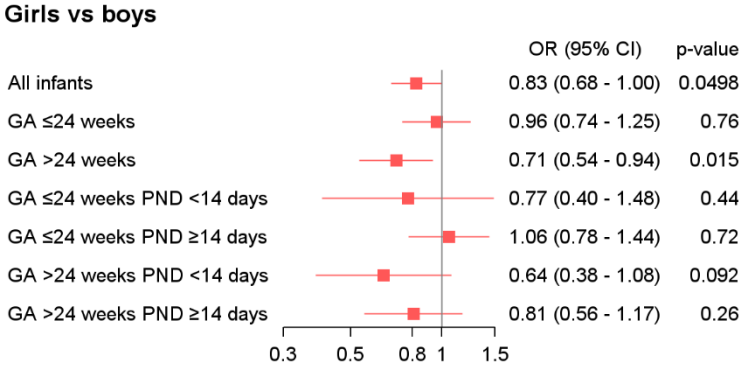
Another interesting finding regarding the progression of ROP was that infants that received parenteral nutrition for ≥14 days showed to have faster progression to ROP treatment, also following adjustment for GA at birth. This suggests that the mechanisms behind the severe medical conditions requiring longer parenteral nutrition further impair eye development, or that longer parenteral nutrition *per se* including its micro-nutrient mixture, and potential undersupply of enteral feeding, are the driving factors.

GENDER-SPECIFIC RISK

Studying sex as risk factor for ROP has come with conflicting results. Some studies have reported that male sex is a significant risk factor while others have reported no differences between the sexes.[79]

In DIGIROP-Birth 1.0 and 2.0, interaction between sex and GA, and sex and PND were significant, meaning that differences between sexes with respect to incidence of ROP treatment differ for different GA and different PND. Following the review of the incidences of ROP treatment per GA, lower risk was observed for girls, but only for higher GA at birth, as shown in ***Figure 23*** below. This suggests that different mechanisms are present and that the type of neonatal care in preventing ROP should consider the infants' sex. Girls with higher GA not requiring prolonged parenteral duration appear to be even more protected against severe ROP. This sex effect is completely diluted among the subgroup of infants with low GA and prolonged parenteral nutrition.

FIGURE 23. RELATION BETWEEN SEX AND SEVERE RETINOPATHY OF PREMATURITY REQUIRING TREATMENT, ADJUSTED FOR GESTATIONAL AGE. *OR = odds-ratio, CI = confidence interval, GA = gestational age, PND = parenteral nutrition duration. OR < 1 means that girls are more protected against severe ROP than boys.*



8.5 ETHICAL CONSIDERATIONS

In studies included in this thesis project no interventions were performed on the infants. An ethical issue to consider is therefore the potential harm to infants by not using DIGIROP prediction models to avoid screening, given the confirmed 100% safety and efficacy of ~50% of the prediction tool on Swedish data. This ethical issue becomes even more sensitive as it involves an increasing and more fragile population, that with time will require a more complex network between different professions at the NICUs. As previously described, ROP screening examinations are on the list of painful neonatal procedures, and the administered dilating eye drops contribute additionally to the impaired well-being of the infant. The true risks and benefits of the models and the reduced number of visits need to be studied in a randomized controlled trial, including long-term follow-up, to confirm or refute the reported benefits of the tool and harms of the ROP screening examinations.

The second ethical issue concerns a situation of a potential failure of the models, which in the worst case could lead to an infant becoming blind. To prevent such situations, it is of critical importance to continue validating the models on temporally different populations and other clinical settings if the models are implemented outside Sweden. In addition, the implementation of the models should be performed in a step-by-step manner, first by sparse examinations, and later by including safety screening examinations at strategically selected time points for the infants discharged by the tool.

9

9 STRENGTHS AND LIMITATIONS

The main strength of the DIGIROP studies is that they are based on a uniquely large cohort worldwide including a complete population of ROP-screened infants in Sweden, at least between years 2008-2020. The input variables are easily obtained by the ophthalmologists from medical records. The DIGIROP models and their decision support tool are available as an online free-of-charge application.

Although registry studies included in this Ph.D. project have the advantage of containing a complete national cohort of ROP-screened infants in Sweden, a limitation is its retrospective design. However, the data were collected using standardized protocols, and extensive efforts were made to validate questioned and missing data points against medical records. Another limitation is the relatively homogenous Swedish population on whom the tool is developed. Although the external validation cohorts originate from two continents, including data from Germany and the US, these cohorts were small and did not represent contemporary data. Further, information about each infant's concurrent medical status covering important risk factors was unavailable. DIGIROP 1.0 is not applicable for infants born at GA<24 weeks due to the lack of Swedish BWSDS reference for these GA.[38] Even though infants with low GA are at high risk for developing ROP and thereby in less need for tools predicting the end of the screening process, DIGIROP 2.0 could include all infants owing to using BW instead of BWSDS. No external validation was performed on the populations from low-income countries with less developed neonatal care and unmonitored oxygen exposure. In these countries the proportion of infants requiring ROP treatment is generally higher. Only limited data on different ethnicities was available for external validation. Continued validation is recommended for different populations and different clinical settings to study further the generalizability and limitations of the developed clinical decision support tool.

10

10 CONCLUSIONS

PAPER I This study showed that PNA, rather than PMA, was a better predictive variable for severe ROP requiring treatment. The hazard function peaked around 12 weeks of PNA and had similar shape for different GA. DIGIROP-Birth 1.0, an individual prediction model for early risk estimation of severe ROP requiring treatment, developed for infants born at GA 24-30 weeks, includes only variables that are easy-obtainable to ophthalmologists and is accessible online. The model appears to be generalizable to temporally different Swedish, German and US data and has shown to have at least as good test or better statistics as other known ROP models.

PAPER II This study developed an individual prediction model, DIGIROP-Screen 1.0, including DIGIROP-Birth 1.0 risk estimates and the status and timing of the first ROP diagnosis, aimed for use during the screening on infants born at GA 24-30 weeks. The model was successfully validated in temporarily and geographically different cohorts. The decision support tool may safely release, at an early stage ~50% of infants that do not need ROP screening, and it has equal or higher sensitivity and specificity than other known ROP models.

PAPER III DIGIROP-Birth 1.0, DIGIROP-Screen 1.0 and their clinical decision support tool show a high predictive ability in a contemporary Swedish cohort. About 50% of infants may be discharged from all ROP screening examinations. All infants, routinely screened, excluding those with clinical indications outside screening criteria were correctly identified as needing ROP screening. Infants with congenital malformation/syndromes, hydrocephalus, and intestinal surgeries should not be discharged from screening by any prediction model.

PAPER IV This study demonstrated that days on parenteral nutrition is a strong predictor of any ROP and severe ROP requiring treatment. Updated DIGIROP-Birth 2.0, DIGIROP-Screen 2.0, and their decision support tool are developed to include all ROP-screened infants. The models and the tool were successfully validated on a temporarily different Swedish cohort. The decision support tool may safely identify low-risk infants that can be released from ROP screening, either early or during the screening. Superiority to other known ROP models is shown.

11

11 FUTURE PERSPECTIVES

11.1 ADAPTATION TO OTHER COUNTRIES

DIGIROP models can be easily adapted to other countries provided that no other important predictors of ROP treatment are superior to GA, weight, sex, and prolonged parenteral nutrition in that setting. On a German data set including 322 infants, DIGIROP models 1.0 achieved 100% sensitivity and 43%-79% specificity at birth and during the screening. Corresponding figures for the US external validation data set, including 366 infants, were 96% sensitivity, incorrectly flagging one infant with a syndrome, and specificity ranging between 41% and 69%. Following the validation of the tool on external data, calibration, and discrimination performance will reveal whether the existing estimates fit the new setting or require re-calibration. [155] To adapt the model to other countries it may be sufficient to update the baseline hazard for the studied outcome, ROP treatment, provided associations between predictors and outcome hold. If necessary, parameter estimates can also be re-estimated. Otherwise, if the discriminative ability and calibration are not satisfied, and a large enough sample is available, a new prediction model can be developed following the same methodology, with or without adding new important predictors.

11.2 IMAGES AND ARTIFICIAL NEURAL NETWORKS

Disagreement between ophthalmological experts the classification of ROP is known.[166 167] Such variabilities introduce errors in the data used for prediction modelling. With the increased use of RetCam (Pleasanton, CA, USA) and storing of wide-field fundus images over the years, several diagnostic tools have been developed for various stages of ROP severity as described in Section 1.3, but also machine learning approaches specifically focusing on reducing the interobserver disagreement error.[168] These tools increase infant safety by considering the timely and correct classification of the disease and improve the input data used for prediction modelling.

Artificial neural network models, a subset of machine learning methodology, are data-driven methods trained to classify and cluster data optimally in multiple

hidden layers against an outcome before an output is provided.[169] The models obtained by these methods cannot evaluate whether knowledge about known risk factors is confirmed. Hence, they are often referred to as the *black box* models. However, if our goal is to obtain an optimal prediction without explaining in what way and how much different variables contribute, these methods are very well suited to solve our problem.

A potential add-on to the DIGIROP models in the future could be to include an early prediction of severe ROP through analysis of vessel morphology characteristics like thickness, tortuosity and growth features using longitudinal fundus images and neural network methodology. A project in cooperation with Dr. Carina Slidsborg and the University of Copenhagen has been initiated for this purpose.

11.3 IMPACT ASSESSMENT AND IMPLEMENTATION SCIENCE

Before DIGIROP clinical decision support tool is implemented in the clinic, an *actual impact assessment* study should be performed, where the tool's actual effect on decision-making, patient outcomes, and healthcare costs are evaluated.[170] Randomized controlled trials, and potentially cluster randomized trials can be planned for this purpose, randomizing some sites to use the prediction tool and others to use usual care. Between-group comparisons assessing clinical usefulness, infant safety, tool's effectiveness, including stress reduction, health economy, and within-group comparisons with respect to pre-trial data, are relevant. Both qualitative and quantitative analyses should be considered.

Implementation science is "*the scientific study of methods to promote the systematic update of research findings and other evidence-based practice into routine practice and, hence, to improve the quality and effectiveness of health services*", meaning that factors defining the *uptake* of the tool into clinical routine are investigated rather than the *health impact* of the tool described above in the *impact assessment* studies.[171] Close discussions with healthcare leaders and healthcare professionals are part of such evaluations.

11.4 PRECISION MEDICINE AND PRECISION HEALTH

The specific focus of this thesis is on individual risk prediction. Modern medicine is developing towards *personalized* and *precision medicine*, which is strongly associated with individual risk predictions. Precision medicine focuses on the *individual* and the individual's optimal treatment and prevention of diseases. Focusing on not one but many individuals, improving *public health* may be achieved, which opens the door to *precision health.* The main goals of *precision health* are to predict, prevent and cure precisely. It is to always be a step ahead and be proactive, with the final goal not just to efficiently treat a disease by early detection, but to completely prevent it, and improve life in the general public.[172]

Placing the fragile population of preterm infants in this context, we should work on prediction models that optimize the treatment of ROP and other severe short- and long-term conditions this population suffers from. We should also work on more personalized preventative treatments to avoid the occurrence of these severe conditions. Why not prevent an infant from being born prematurely and erase the complications that prematurity causes? Ultimately, the key lies in the mother's health, her lifestyle, and the social and physical environment that surrounds her.

12

12 ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the people that have been part of this thesis project and supported me in different ways over the years.

Prof. Ann Hellström, my main supervisor, thank you for giving me this opportunity to develop, for your unwavering availability, despite your demanding schedule. I am deeply grateful for the knowledge you have shared with me, and for your wise contributions to the papers and thesis. Your encouragement has been crucial in helping me achieve this goal, and I deeply appreciate all that you have done for me. I admire you greatly, both as a person and as a professional - you are a true role model for me to follow.

Prof. Staffan Nilsson, my co-supervisor, thank you for introducing me to the world of education, which helped calm my nervousness. Your statistical contributions in the papers and thesis have been instrumental. I admire your exceptional pedagogical qualities, which have inspired me to become a better statistician and researcher.

Dr. Helena Johansson, my co-supervisor, thank you for your availability and support throughout my academic journey. Your valuable contributions and insightful discussions on statistical modelling have been very important, and I am truly grateful for your input in the papers and thesis.

Sr. Prof. Kerstin Albertsson-Wikland, my co-supervisor, thank you for your pivotal role in helping me take this academic path, by posing that question (again) almost five years ago. Thank you for your encouragement throughout the challenging periods and your guidance in structuring my working life. Our engaging discussions on my projects, this thesis and research in general, have been greatly appreciated. You are a woman of steel with the biggest heart, and I am deeply grateful to have had you on this journey.

Prof. Hans Wedel, to whom I would like to express my sincere appreciation for being a mentor to me throughout this journey. I always learn something new from you, be it about statistics, research, history, or life in general. Our cooperation on different projects means very much to me. I am constantly amazed by your professional experience and wisdom, and I hope that we will continue to work together for many years to come.

Prof. Lois E.H. Smith, **Dr. Anna-Lena Hård**, **Ass. Prof. Chatarina Löfqvist**, **Ass. Prof. Pia Lundgren**, my co-authors and members of the closest working group together with my supervisors, deepest thank you for your invaluable contributions to the papers and thesis, as well as for all the insightful discussions and swift reviews, even on late nights and weekends. I cannot thank you enough for your support and dedication.

Prof. Gerd Holmström, the register holder of SWEDROP and my co-author, deepest thank you for your expertise, support with the data, and important discussions in the projects. The prediction models would not have been possible without the input from this unique and high-quality register. To my co-authors and SWEDROP members **Dr. Lotta Gränse**, **Dr. Eva Larsson**, **Dr. Abbas Al-Hawasi**, **Dr. Birgitta Sunnqvist**, **Dr. Kristina Tornqvist**, and **Dr. Agneta Wallin**, I want to express my sincere appreciation for all your contributions to the papers.

Dr. Andreas Stahl, Dr. Marie-Christine Bründer, **Dr. Wolf A. Lagrèze**, **Dr. Carolyn Wu**, **Dr. Bradley Yoder**, and **Prof. Elizabeth Hartnett**, my co-authors, thank you for the contributions to the external validation of the DIGIROP prediction models, using your international data. Your valuable input in the papers has been greatly appreciated.

Carola Pfeiffer-Mosesson, **Lena Kjellberg**, **Ulrika Sjöbom**, **Dr. Margareta Höök-Wikstrand**, **Ann-Charlotte Andersson**, my research colleagues, thank you for your excellent help with the validation of the SWEDROP data, and WINROP calculations.

Dr. Karin Sävman, **Ass. Prof. Ingrid Hansen-Pupp**, and **Dr. Dirk Wackernagel**, thank you for sharing your knowledge on neonatology and nutritional strategies.

Prof. Emeritus Anders Odén, an exceptional statistician who developed the Poisson BASIC macro, and **Artemis Mårtensson**, my former colleague who transcribed the BASIC macro into SAS, I am deeply grateful for everything you have taught me. To **Nils-Gunnar Pehrsson**, my former employer, and **colleagues at Statistiska konsultgruppen**, I am thankful for the highly educational environment over the years and for granting me permission to use the Poisson SAS macro in all my PhD projects.

To **Dr. William Hellström**, **Ass. Prof. Anders Nilsson**, **Ass. Prof. Gunnel Hellgren**, my research colleagues, thank you for cheering and supporting me along the way.

My dear friends since forever, **Mille**, **Heidi**, and **Danijela**, I want to express my sincere gratitude for the wonderful memories we have shared, and for your continued presence in my life.

My dear mother- and father-in-law **Eva** and **Bo**, as well as sister- and brother-in-law **Suzanne** and **Mats**, I want to express my deep appreciation for your kindness, love, and support. Your presence in our lives has been invaluable, and we are truly grateful for all that you do for us. Special thanks to Suzanne for creating the party invitations.

Losing Bosse earlier this year was a great loss to us all, and we will always remember him with love and fondness.

My dear **Almira**, **Aida** and **Almir**, thank you for being not just my siblings, but also my friends for life. Our parents and grandparents succeeded in fostering strong bonds between us, and I am so grateful for that. I feel truly fortunate to have you and your families in our lives. Special thanks to Aida for painting *The Iris*.

My dear mother **Jasmina** and father **Faruk**, my deepest thank you for always being there for me, and for your unconditional love and generosity towards me, Philip, and our boys.

My dearest **Philip**, thank you for your love, and for being my determined support throughout this journey and life. You, **Edvin** and **Miro** mean everything to me.

This thesis was financed by the Sahlgrenska Academy, and Prof. Ann Hellström's grants: the Swedish Medical Research Council, The Gothenburg Medical Society and Government grants under the ALF agreement, Knut and Alice Wallenberg Clinical Scholars, and De Blindas Vänner.

REFERENCES

1. Engle WA, American Academy of Pediatrics Committee on F, Newborn. Age terminology during the perinatal period. *Pediatrics* 2004;114(5):1362-4. doi: 10.1542/peds.2004-1915
2. Office on Women's Health - Stages on Pregnancy [Available from: https://www.womenshealth.gov/pregnancy/youre-pregnant-now-what/stages-pregnancy accessed April 11 2023 2023.
3. Heidary G, Vanderveen D, Smith LE. Retinopathy of prematurity: current concepts in molecular pathogenesis. *Semin Ophthalmol* 2009;24(2):77-81. doi: 10.1080/08820530902800314
4. World Health Organization - Preterm birth [Available from: https://www.who.int/news-room/fact-sheets/detail/preterm-birth accessed April 11 2023 2023.
5. Blencowe H, Cousens S, Oestergaard MZ, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet* 2012;379(9832):2162-72. doi: 10.1016/S0140-6736(12)60820-4
6. Organization WH. Born Too Soon - The Global Action Report on Preterm Birth, 2012.
7. Lawn JE, Blencowe H, Oza S, et al. Every Newborn: progress, priorities, and potential beyond survival. *Lancet* 2014;384(9938):189-205. doi: 10.1016/S0140-6736(14)60496-7 [published Online First: 20140519]
8. Swedish Neonatal Quality Register - Yearly Report 2021 2021 [Available from: https://www.medscinet.com/PNQ/uploads/website/Neonatalv%C3%A5rdsregistrets%20%C3%85rsrapport%202021%20(final%20version).pdf accessed April 11 2023 2023.
9. Swedish Pregnancy Register - Yearly Report 2021 2021 [Available from: https://www.medscinet.com/GR/uploads/hemsida/dokumentarkiv/GR%20%C3%85rsrapport%202021_3.0.pdf accessed April 11 2023 2023.
10. Norman M, Hallberg B, Abrahamsson T, et al. Association Between Year of Birth and 1-Year Survival Among Extremely Preterm Infants in Sweden During 2004-2007 and 2014-2016. *JAMA* 2019;321(12):1188-99. doi: 10.1001/jama.2019.2021
11. Domellof M, Jonsson B. The Swedish Approach to Management of Extreme Prematurity at the Borderline of Viability: A Historical and Ethical Perspective. *Pediatrics* 2018;142(Suppl 1):S533-S38. doi: 10.1542/peds.2018-0478C
12. Lundgren P, Morsing E, Hard AL, et al. National cohort of infants born before 24 gestational weeks showed increased survival rates but no improvement in neonatal morbidity. *Acta Paediatr* 2022;111(8):1515-25. doi: 10.1111/apa.16354 [published Online First: 20220412]
13. Mayo Clinic - Premature Birth [Available from: https://www.mayoclinic.org/diseases-conditions/premature-birth/symptoms-causes/syc-20376730 accessed April 11 2023 2023.
14. Costeloe KL, Hennessy EM, Haider S, et al. Short term outcomes after extreme preterm birth in England: comparison of two birth cohorts in 1995 and 2006 (the EPICure studies). *BMJ* 2012;345:e7976. doi: 10.1136/bmj.e7976 [published Online First: 20121204]
15. Group E. Incidence of and risk factors for neonatal morbidity after active perinatal care: extremely preterm infants study in Sweden (EXPRESS). *Acta Paediatr* 2010;99(7):978-92. doi: 10.1111/j.1651-2227.2010.01846.x [published Online First: 20100426]

16. Lundgren P, Jacobson L, Hard AL, et al. High rate and large intercentre variability in retreatment of retinopathy of prematurity in infants born <24 gestational weeks. *BMJ Open Ophthalmol* 2021;6(1):e000695. doi: 10.1136/bmjophth-2020-000695 [published Online First: 20210421]
17. Saigal S, Doyle LW. An overview of mortality and sequelae of preterm birth from infancy to adulthood. *Lancet* 2008;371(9608):261-9. doi: 10.1016/S0140-6736(08)60136-1
18. Pravia CI, Benny M. Long-term consequences of prematurity. *Cleve Clin J Med* 2020;87(12):759-67. doi: 10.3949/ccjm.87a.19108 [published Online First: 20201123]
19. Hellstrom A, Jacobson L, Al-Hawasi A, et al. Retrospective evaluation of ophthalmological and neurological outcomes for infants born before 24 weeks gestational age in a Swedish cohort. *BMJ Open* 2022;12(8):e055567. doi: 10.1136/bmjopen-2021-055567 [published Online First: 20220803]
20. World Health Organization - Exclusive breastfeeding for six months best for babies everywhere [Available from: https://apps.who.int/mediacentre/news/statements/2011/breastfeeding_20110115/en/index.html#:~:text=Exclusive%20breastfeeding%20for%20six%20months%20best%20for%20babies,to%20the%20age%20of%20two%20years%20or%20beyond. accessed April 11 2023 2023.
21. Meek JY, Noble L, Section on B. Policy Statement: Breastfeeding and the Use of Human Milk. *Pediatrics* 2022;150(1) doi: 10.1542/peds.2022-057988
22. Embleton ND, Moltu SJ, Lapillonne A, et al. Enteral Nutrition in Preterm Infants (2022): A Position Paper From the ESPGHAN Committee on Nutrition and Invited Experts. *J Pediatr Gastroenterol Nutr* 2023;76(2):248-68. doi: 10.1097/MPG.0000000000003642 [published Online First: 20221028]
23. Hoban R, Bowker RM, Gross ME, et al. Maternal production of milk for infants in the neonatal intensive care unit. *Semin Perinatol* 2021;45(2):151381. doi: 10.1016/j.semperi.2020.151381 [published Online First: 20210102]
24. Stephens BE, Walden RV, Gargus RA, et al. First-week protein and energy intakes are associated with 18-month developmental outcomes in extremely low birth weight infants. *Pediatrics* 2009;123(5):1337-43. doi: 10.1542/peds.2008-0211
25. Stoltz Sjostrom E, Ohlund I, Ahlsson F, et al. Nutrient intakes independently affect growth in extremely preterm infants: results from a population-based study. *Acta Paediatr* 2013;102(11):1067-74. doi: 10.1111/apa.12359 [published Online First: 20130806]
26. Socialstyrelsen - Vård av extremt för tidigt födda barn [Available from: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/vagledning/2014-9-10.pdf#:~:text=Att%20f%C3%A5%20ett%20extremt%20f%C3%B6r%20tidigt%20f%C3%B6tt%20barn,och%20kontinuerligt%20informeras%20om%20barnets%20tillst%C3%A5nd%20och%20prognos. accessed April 11 2023 2023.
27. Mihatsch WA, Braegger C, Bronsky J, et al. ESPGHAN/ESPEN/ESPR/CSPEN guidelines on pediatric parenteral nutrition. *Clin Nutr* 2018;37(6 Pt B):2303-05. doi: 10.1016/j.clnu.2018.05.029 [published Online First: 20180607]
28. Moltu SJ, Bronsky J, Embleton N, et al. Nutritional Management of the Critically Ill Neonate: A Position Paper of the ESPGHAN Committee on Nutrition. *J Pediatr Gastroenterol Nutr* 2021;73(2):274-89. doi: 10.1097/MPG.0000000000003076
29. De Nardo MC, Mario CD, Laccetta G, et al. Enteral and parenteral energy intake and neurodevelopment in preterm infants: A systematic review. *Nutrition*

2022;97:111572. doi: 10.1016/j.nut.2021.111572 [published Online First: 20211222]
30. Gao L, Shen W, Wu F, et al. Effect of early initiation of enteral nutrition on short-term clinical outcomes of very premature infants: A national multicenter cohort study in China. *Nutrition* 2022;107:111912. doi: 10.1016/j.nut.2022.111912 [published Online First: 20221112]
31. Hellstrom A, Engstrom E, Hard AL, et al. Postnatal serum insulin-like growth factor I deficiency is associated with retinopathy of prematurity and other complications of premature birth. *Pediatrics* 2003;112(5):1016-20. doi: 10.1542/peds.112.5.1016
32. Hellstrom A, Smith LE, Dammann O. Retinopathy of prematurity. *Lancet* 2013;382(9902):1445-57. doi: 10.1016/S0140-6736(13)60178-6 [published Online First: 20130617]
33. Hard AL, Smith LE, Hellstrom A. Nutrition, insulin-like growth factor-1 and retinopathy of prematurity. *Semin Fetal Neonatal Med* 2013;18(3):136-42. doi: 10.1016/j.siny.2013.01.006 [published Online First: 20130218]
34. Anvekar A, Athikarisamy S, Rao S, et al. Time to regain birth weight - a marker to predict the severity of retinopathy of prematurity? *BMC Pediatr* 2021;21(1):540. doi: 10.1186/s12887-021-03027-x [published Online First: 20211202]
35. Klevebro S, Lundgren P, Hammar U, et al. Cohort study of growth patterns by gestational age in preterm infants developing morbidity. *BMJ Open* 2016;6(11):e012872. doi: 10.1136/bmjopen-2016-012872 [published Online First: 20161117]
36. Fenton TR, Kim JH. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. *BMC Pediatr* 2013;13:59. doi: 10.1186/1471-2431-13-59 [published Online First: 20130420]
37. Fenton Preterm Growth Charts [Available from: https://ucalgary.ca/resource/preterm-growth-chart/preterm-growth-chart accessed April 11 2023 2023.
38. Niklasson A, Albertsson-Wikland K. Continuous growth reference from 24th week of gestation to 24 months by gender. *BMC Pediatr* 2008;8:8. doi: 10.1186/1471-2431-8-8 [published Online First: 20080229]
39. Terry TL. Extreme Prematurity and Fibroblastic Overgrowth of Persistent Vascular Sheath Behind Each Crystalline Lens: I. Preliminary report. *Am J Ophthalmol* 2018;192:xxviii. doi: 10.1016/j.ajo.2018.05.024
40. Ashton N, Ward B, Serpell G. Effect of oxygen on developing retinal vessels with particular reference to the problem of retrolental fibroplasia. *Br J Ophthalmol* 1954;38(7):397-432. doi: 10.1136/bjo.38.7.397
41. Shweiki D, Itin A, Soffer D, et al. Vascular endothelial growth factor induced by hypoxia may mediate hypoxia-initiated angiogenesis. *Nature* 1992;359(6398):843-5. doi: 10.1038/359843a0
42. Smith LE. Through the eyes of a child: understanding retinopathy through ROP the Friedenwald lecture. *Invest Ophthalmol Vis Sci* 2008;49(12):5177-82. doi: 10.1167/iovs.08-2584 [published Online First: 20080815]
43. Hellstrom A, Carlsson B, Niklasson A, et al. IGF-I is critical for normal vascularization of the human retina. *J Clin Endocrinol Metab* 2002;87(7):3413-6. doi: 10.1210/jcem.87.7.8629
44. Hellstrom A, Perruzzi C, Ju M, et al. Low IGF-I suppresses VEGF-survival signaling in retinal endothelial cells: direct correlation with clinical retinopathy of prematurity. *Proc Natl Acad Sci U S A* 2001;98(10):5804-8. doi: 10.1073/pnas.101113998 [published Online First: 20010501]

45. Aiello LP, Pierce EA, Foley ED, et al. Suppression of retinal neovascularization in vivo by inhibition of vascular endothelial growth factor (VEGF) using soluble VEGF-receptor chimeric proteins. *Proc Natl Acad Sci U S A* 1995;92(23):10457-61. doi: 10.1073/pnas.92.23.10457
46. Blencowe H, Lawn JE, Vazquez T, et al. Preterm-associated visual impairment and estimates of retinopathy of prematurity at regional and global levels for 2010. *Pediatr Res* 2013;74 Suppl 1:35-49. doi: 10.1038/pr.2013.205
47. Gilbert C, Fielder A, Gordillo L, et al. Characteristics of infants with severe retinopathy of prematurity in countries with low, moderate, and high levels of development: implications for screening programs. *Pediatrics* 2005;115(5):e518-25. doi: 10.1542/peds.2004-1180 [published Online First: 20050401]
48. Holmstrom G, Hellstrom A, Granse L, et al. New modifications of Swedish ROP guidelines based on 10-year data from the SWEDROP register. *Br J Ophthalmol* 2020;104(7):943-49. doi: 10.1136/bjophthalmol-2019-314874 [published Online First: 20191101]
49. Norman M, Hellstrom A, Hallberg B, et al. Prevalence of Severe Visual Disability Among Preterm Children With Retinopathy of Prematurity and Association With Adherence to Best Practice Guidelines. *JAMA Netw Open* 2019;2(1):e186801. doi: 10.1001/jamanetworkopen.2018.6801 [published Online First: 20190104]
50. Chiang MF, Quinn GE, Fielder AR, et al. International Classification of Retinopathy of Prematurity, Third Edition. *Ophthalmology* 2021;128(10):e51-e68. doi: 10.1016/j.ophtha.2021.05.031 [published Online First: 20210708]
51. Fierson WM, Ophthalmology AAOPSo, American Academy Of O, et al. Screening Examination of Premature Infants for Retinopathy of Prematurity. *Pediatrics* 2018;142(6) doi: 10.1542/peds.2018-3061
52. International Committee for the Classification of Retinopathy of P. The International Classification of Retinopathy of Prematurity revisited. *Arch Ophthalmol* 2005;123(7):991-9. doi: 10.1001/archopht.123.7.991
53. Smith LEH, Hellstrom A, Stahl A, et al. Development of a Retinopathy of Prematurity Activity Scale and Clinical Outcome Measures for Use in Clinical Trials. *JAMA Ophthalmol* 2019;137(3):305-11. doi: 10.1001/jamaophthalmol.2018.5984
54. Pivodic A, Nilsson S, Stahl A, et al. Validation of the Retinopathy of Prematurity Activity Scale (ROP-ActS) using retrospective clinical data. *Acta Ophthalmol* 2021;99(2):201-06. doi: 10.1111/aos.14532 [published Online First: 20200626]
55. Pivodic A, Johansson H, Smith LE, et al. Evaluation of the Retinopathy of Prematurity Activity Scale (ROP-ActS) in a randomised controlled trial aiming for prevention of severe ROP: a substudy of the Mega Donna Mega trial. *BMJ Open Ophthalmol* 2022;7(1):e000923. doi: 10.1136/bmjophth-2021-000923 [published Online First: 20220408]
56. Gupta K, Campbell JP, Taylor S, et al. A Quantitative Severity Scale for Retinopathy of Prematurity Using Deep Learning to Monitor Disease Regression After Treatment. *JAMA Ophthalmol* 2019;137(9):1029-36. doi: 10.1001/jamaophthalmol.2019.2442 [published Online First: 20190703]
57. Mora JS, Waite C, Gilbert CE, et al. A worldwide survey of retinopathy of prematurity screening. *Br J Ophthalmol* 2018;102(1):9-13. doi: 10.1136/bjophthalmol-2017-310709 [published Online First: 20170830]
58. Wilkinson AR, Adams GGW, Fleck BW, et al. UK screening and treatment of retinopathy of prematurity Updated 2022 Guidelines. *Early Hum Dev* 2023;177-178:105715. doi: 10.1016/j.earlhumdev.2023.105715 [published Online First: 20230125]

59. Retinopathy of Prematurity Screening Guidelines in Germany [Available from: https://register.awmf.org/assets/guidelines/024-010l_S2k_Augenaerztliche_Screening-Untersuchung_Fr%C3%BChgeborene_2020-07.pdf accessed March 17, 2023 2023.
60. Barjol A, Lux AL, Dureau P, et al. Evaluation and modification of French screening guidelines for retinopathy of prematurity. *Acta Ophthalmol* 2022;100(7):e1451-e54. doi: 10.1111/aos.15091 [published Online First: 20220126]
61. Ferrer Novella C, Gonzalez Viejo I, Pueyo Royo V, et al. Screening program for retinopathy of prematurity in Spain. *Arch Soc Esp Oftalmol* 2013;88(5):184-8. doi: 10.1016/j.oftal.2012.07.030 [published Online First: 20121018]
62. Retinopathy of Prematurity Screening Guidelines in Portugal [Available from: https://www.spneonatologia.pt/wp-content/uploads/2016/11/2014-ROP.pdf accessed March 17, 2023 2023.
63. Mataftsi A, Moutzouri S, Karagianni P, et al. Retinopathy of prematurity occurrence and evaluation of screening policy in a large tertiary Greek cohort. *Int Ophthalmol* 2020;40(2):385-91. doi: 10.1007/s10792-019-01195-5 [published Online First: 20191022]
64. Ministry of Health & Family Welfare India. Guidelines for Universal Eye Screening in Newborns Including Retinopathy of Prematurity [Available from: http://www.nhm.gov.in/images/pdf/programmes/RBSK/Resource_Documents/Revised_ROP_Guidelines-Web_Optimized.pdf accessed April 11 2023 2023.
65. Swedish national guidelines for screening and management of ROP 2021 [Available from: https://www.medscinet.com/ROP/uploads/Nationella%20guidelines%20ROP-%20Revidering%2016.12.2022.pdf accessed April 11 2023 2023.
66. Holmstrom GE, Hellstrom A, Jakobsson PG, et al. Swedish national register for retinopathy of prematurity (SWEDROP) and the evaluation of screening in Sweden. *Arch Ophthalmol* 2012;130(11):1418-24. doi: 10.1001/archophthalmol.2012.2357
67. Swedish national guidelines for screening and management of ROP 2012 [Available from: https://neo.barnlakarforeningen.se/wp-content/uploads/sites/14/2014/03/ROP_2012.pdf accessed April 11 2023 2023.
68. Swedish National Registry for Retinopathy of Prematurity [Available from: https://www.medscinet.com/ROP/ accessed April 11 2023 2023.
69. Swedish Neonatal Quality Register [Available from: https://www.medscinet.com/pnq/default.aspx accessed April 11 2023 2023.
70. Chiang MF, Wang L, Busuioc M, et al. Telemedical retinopathy of prematurity diagnosis: accuracy, reliability, and image quality. *Arch Ophthalmol* 2007;125(11):1531-8. doi: 10.1001/archopht.125.11.1531
71. Wang J, Ju R, Chen Y, et al. Automated retinopathy of prematurity screening using deep neural networks. *EBioMedicine* 2018;35:361-68. doi: 10.1016/j.ebiom.2018.08.033 [published Online First: 20180827]
72. Brown JM, Campbell JP, Beers A, et al. Automated Diagnosis of Plus Disease in Retinopathy of Prematurity Using Deep Convolutional Neural Networks. *JAMA Ophthalmol* 2018;136(7):803-10. doi: 10.1001/jamaophthalmol.2018.1934
73. Taylor S, Brown JM, Gupta K, et al. Monitoring Disease Progression With a Quantitative Severity Scale for Retinopathy of Prematurity Using Deep Learning. *JAMA Ophthalmol* 2019;137(9):1022-8. doi: 10.1001/jamaophthalmol.2019.2433 [published Online First: 20190703]

74. Wang SK, Korot E, Zaidi M, et al. Modeling absolute zone size in retinopathy of prematurity in relation to axial length. *Sci Rep* 2022;12(1):4717. doi: 10.1038/s41598-022-08680-5 [published Online First: 20220318]
75. Huang YP, Basanta H, Kang EY, et al. Automated detection of early-stage ROP using a deep convolutional neural network. *Br J Ophthalmol* 2021;105(8):1099-103. doi: 10.1136/bjophthalmol-2020-316526 [published Online First: 20200823]
76. Ataer-Cansizoglu E, Bolon-Canedo V, Campbell JP, et al. Computer-Based Image Analysis for Plus Disease Diagnosis in Retinopathy of Prematurity: Performance of the "i-ROP" System and Image Features Associated With Expert Diagnosis. *Transl Vis Sci Technol* 2015;4(6):5. doi: 10.1167/tvst.4.6.5 [published Online First: 20151130]
77. Hu J, Chen Y, Zhong J, et al. Automated Analysis for Retinopathy of Prematurity by Deep Neural Networks. *IEEE Trans Med Imaging* 2019;38(1):269-79. doi: 10.1109/TMI.2018.2863562 [published Online First: 20180806]
78. Campbell JP, Kim SJ, Brown JM, et al. Evaluation of a Deep Learning-Derived Quantitative Retinopathy of Prematurity Severity Scale. *Ophthalmology* 2021;128(7):1070-76. doi: 10.1016/j.ophtha.2020.10.025 [published Online First: 20201027]
79. Kim SJ, Port AD, Swan R, et al. Retinopathy of prematurity: a review of risk factors and their clinical significance. *Surv Ophthalmol* 2018;63(5):618-37. doi: 10.1016/j.survophthal.2018.04.002 [published Online First: 20180419]
80. Gunn TR, Aranda JV, Little J. Incidence of retrolental fibroplasia. *Lancet* 1978;1(8057):216-7. doi: 10.1016/s0140-6736(78)90660-8
81. Shohat M, Reisner SH, Krikler R, et al. Retinopathy of prematurity: incidence and risk factors. *Pediatrics* 1983;72(2):159-63.
82. Vanhaesebrouck S, Vanhole C, de Zegher F, et al. Influence of duration of parenteral nutrition on retinopathy of prematurity. *Arch Dis Child Fetal Neonatal Ed* 2008;93(2):F170. doi: 10.1136/adc.2007.128991
83. Niwald A, Piotrowski A, Gralek M. [Analysis of some of the possible neonatal risk factors of development of retinopathy of prematurity]. *Klin Oczna* 2008;110(1-3):31-4.
84. Petrachkova MS, Saidasheva EI, Petrachkov DV, et al. [Modern approaches to predicting the development of active type 1 retinopathy of prematurity]. *Vestn Oftalmol* 2019;135(4):50-59. doi: 10.17116/oftalma201913504150
85. Bassiouny MR. Risk factors associated with retinopathy of prematurity: a study from Oman. *J Trop Pediatr* 1996;42(6):355-8. doi: 10.1093/tropej/42.6.355
86. Porcelli PJ, Weaver RG, Jr. The influence of early postnatal nutrition on retinopathy of prematurity in extremely low birth weight infants. *Early Hum Dev* 2010;86(6):391-6. doi: 10.1016/j.earlhumdev.2010.05.015 [published Online First: 20100618]
87. Ali AA, Gomaa NAS, Awadein AR, et al. Retrospective cohort study shows that the risks for retinopathy of prematurity included birth age and weight, medical conditions and treatment. *Acta Paediatr* 2017;106(12):1919-27. doi: 10.1111/apa.14019 [published Online First: 20170918]
88. Hellstrom W, Martinsson T, Morsing E, et al. Low fraction of fetal haemoglobin is associated with retinopathy of prematurity in the very preterm infant. *Br J Ophthalmol* 2022;106(7):970-74. doi: 10.1136/bjophthalmol-2020-318293 [published Online First: 20210205]
89. Hellstrom A, Nilsson AK, Wackernagel D, et al. Effect of Enteral Lipid Supplement on Severe Retinopathy of Prematurity: A Randomized Clinical Trial. *JAMA Pediatr* 2021;175(4):359-67. doi: 10.1001/jamapediatrics.2020.5653
90. Wendel K, Aas MF, Gunnarsdottir G, et al. Effect of arachidonic and docosahexaenoic acid supplementation on respiratory outcomes and neonatal morbidities in preterm

infants. *Clin Nutr* 2023;42(1):22-28. doi: 10.1016/j.clnu.2022.11.012 [published Online First: 20221117]
91. Hellstrom A, Pivodic A, Granse L, et al. Association of Docosahexaenoic Acid and Arachidonic Acid Serum Levels With Retinopathy of Prematurity in Preterm Infants. *JAMA Netw Open* 2021;4(10):e2128771. doi: 10.1001/jamanetworkopen.2021.28771 [published Online First: 20211001]
92. Early Treatment For Retinopathy Of Prematurity Cooperative G. Revised indications for the treatment of retinopathy of prematurity: results of the early treatment for retinopathy of prematurity randomized trial. *Arch Ophthalmol* 2003;121(12):1684-94. doi: 10.1001/archopht.121.12.1684
93. Good WV, Early Treatment for Retinopathy of Prematurity Cooperative G. Final results of the Early Treatment for Retinopathy of Prematurity (ETROP) randomized trial. *Trans Am Ophthalmol Soc* 2004;102:233-48; discussion 48-50.
94. Dammann O, Hartnett ME, Stahl A. Retinopathy of prematurity. *Dev Med Child Neurol* 2022 doi: 10.1111/dmcn.15468 [published Online First: 20221121]
95. Chang E, Josan AS, Purohit R, et al. A Network Meta-Analysis of Retreatment Rates following Bevacizumab, Ranibizumab, Aflibercept, and Laser for Retinopathy of Prematurity. *Ophthalmology* 2022;129(12):1389-401. doi: 10.1016/j.ophtha.2022.06.042 [published Online First: 20220714]
96. Steyerberg EW. Clinical Prediction Models - A Practical Approach to Development, Validation, and Updating. New York: Springer 2009.
97. Binuya MAE, Engelhardt EG, Schats W, et al. Methodological guidance for the evaluation and updating of clinical prediction models: a systematic review. *BMC Med Res Methodol* 2022;22(1):316. doi: 10.1186/s12874-022-01801-8 [published Online First: 20221212]
98. Lofqvist C, Andersson E, Sigurdsson J, et al. Longitudinal postnatal weight and insulin-like growth factor I measurements in the prediction of retinopathy of prematurity. *Arch Ophthalmol* 2006;124(12):1711-8. doi: 10.1001/archopht.124.12.1711
99. Hellstrom A, Hard AL, Engstrom E, et al. Early weight gain predicts retinopathy in preterm infants: new, simple, efficient approach to screening. *Pediatrics* 2009;123(4):e638-45. doi: 10.1542/peds.2008-2697 [published Online First: 20090316]
100. WINROP Online Monitoring System [Available from: https://winrop.com/ accessed April 11 2023 2023.
101. Binenbaum G, Bell EF, Donohue P, et al. Development of Modified Screening Criteria for Retinopathy of Prematurity: Primary Results From the Postnatal Growth and Retinopathy of Prematurity Study. *JAMA Ophthalmol* 2018;136(9):1034-40. doi: 10.1001/jamaophthalmol.2018.2753
102. Yang MB, Donovan EF. Risk analysis and an alternative protocol for reduction of screening for retinopathy of prematurity. *J AAPOS* 2009;13(6):539-45. doi: 10.1016/j.jaapos.2009.10.002
103. Slidsborg C, Forman JL, Rasmussen S, et al. A new risk-based screening criterion for treatment-demanding retinopathy of prematurity in Denmark. *Pediatrics* 2011;127(3):e598-606. doi: 10.1542/peds.2010-1974 [published Online First: 20110214]
104. Iu LPL, Yip WWK, Lok JYC, et al. Prediction model to predict type 1 retinopathy of prematurity using gestational age and birth weight (PW-ROP). *Br J Ophthalmol* 2022 doi: 10.1136/bjophthalmol-2021-320670 [published Online First: 20220217]

105. Hardy RJ, Palmer EA, Dobson V, et al. Risk analysis of prethreshold retinopathy of prematurity. *Arch Ophthalmol* 2003;121(12):1697-701. doi: 10.1001/archopht.121.12.1697
106. Binenbaum G, Ying GS, Quinn GE, et al. A clinical prediction model to stratify retinopathy of prematurity risk using postnatal weight gain. *Pediatrics* 2011;127(3):e607-14. doi: 10.1542/peds.2010-2240 [published Online First: 20110214]
107. Binenbaum G, Ying G-s, Quinn GE, et al. The CHOP Postnatal Weight Gain, Birth Weight, and Gestational Age Retinopathy of Prematurity Risk Model. *Archives of Ophthalmology* 2012;130(12):1560-65.
108. Eckert GU, Fortes Filho JB, Maia M, et al. A predictive score for retinopathy of prematurity in very low birth weight preterm infants. *Eye (Lond)* 2012;26(3):400-6. doi: 10.1038/eye.2011.334 [published Online First: 20111223]
109. Ying GS, Quinn GE, Wade KC, et al. Predictors for the development of referral-warranted retinopathy of prematurity in the telemedicine approaches to evaluating acute-phase retinopathy of prematurity (e-ROP) study. *JAMA Ophthalmol* 2015;133(3):304-11. doi: 10.1001/jamaophthalmol.2014.5185
110. Cao JH, Wagner BD, Cerda A, et al. Colorado retinopathy of prematurity model: a multi-institutional validation study. *J AAPOS* 2016;20(3):220-5. doi: 10.1016/j.jaapos.2016.01.017 [published Online First: 20160507]
111. Slidsborg C, Jensen A, Forman JL, et al. Neonatal Risk Factors for Treatment-Demanding Retinopathy of Prematurity: A Danish National Study. *Ophthalmology* 2016;123(4):796-803. doi: 10.1016/j.ophtha.2015.12.019 [published Online First: 20160204]
112. Ricard CA, Dammann CEL, Dammann O. Screening Tool for Early Postnatal Prediction of Retinopathy of Prematurity in Preterm Newborns (STEP-ROP). *Neonatology* 2017;112(2):130-36. doi: 10.1159/000464459 [published Online First: 20170513]
113. Ludwig CA, Greven MA, Moshfeghi DM. Predictors of treatment-warranted retinopathy of prematurity in the SUNDROP cohort: influence of photographic features. *Graefes Arch Clin Exp Ophthalmol* 2017;255(10):1935-46. doi: 10.1007/s00417-017-3745-3 [published Online First: 20170807]
114. Binenbaum G, Ying GS, Tomlinson LA, et al. Validation of the Children's Hospital of Philadelphia Retinopathy of Prematurity (CHOP ROP) Model. *JAMA Ophthalmol* 2017;135(8):871-77. doi: 10.1001/jamaophthalmol.2017.2295
115. Gurwin J, Tomlinson LA, Quinn GE, et al. A Tiered Approach to Retinopathy of Prematurity Screening (TARP) Using a Weight Gain Predictive Model and a Telemedicine System. *JAMA Ophthalmol* 2017;135(2):131-36. doi: 10.1001/jamaophthalmol.2016.5203
116. McCauley K, Chundu A, Song H, et al. Implementation of a Clinical Prediction Model Using Daily Postnatal Weight Gain, Birth Weight, and Gestational Age to Risk Stratify ROP. *J Pediatr Ophthalmol Strabismus* 2018;55(5):326-34. doi: 10.3928/01913913-20180405-02 [published Online First: 20180619]
117. Ahmed IS, Badeeb AA. The Alexandria retinopathy of prematurity model (Alex-ROP): postnatal weight gain screening algorithm application in a developing country. *Int J Ophthalmol* 2019;12(2):296-301. doi: 10.18240/ijo.2019.02.18 [published Online First: 20190218]
118. Ying GS, Bell EF, Donohue P, et al. Perinatal Risk Factors for the Retinopathy of Prematurity in Postnatal Growth and Rop Study. *Ophthalmic Epidemiol* 2019;26(4):270-78. doi: 10.1080/09286586.2019.1606259 [published Online First: 20190423]

119. Cheng QE, Daniel E, Pan W, et al. Plus Disease in Telemedicine Approaches to Evaluating Acute-Phase ROP (e-ROP) Study: Characteristics, Predictors, and Accuracy of Image Grading. *Ophthalmology* 2019;126(6):868-75. doi: 10.1016/j.ophtha.2019.01.021 [published Online First: 20190125]
120. Estrada MM, Tomlinson LA, Yu Y, et al. Daily Oxygen Supplementation and Risk of Retinopathy of Prematurity. *Ophthalmic Epidemiol* 2022:1-9. doi: 10.1080/09286586.2022.2111687 [published Online First: 20220912]
121. Siswanto JE, Adisasmita AC, Ronoatmodjo S, et al. A risk scoring model to predict progression of retinopathy of prematurity for Indonesia. *PLoS One* 2023;18(2):e0281284. doi: 10.1371/journal.pone.0281284 [published Online First: 20230203]
122. Coyner AS, Chen JS, Singh P, et al. Single-Examination Risk Prediction of Severe Retinopathy of Prematurity. *Pediatrics* 2021;148(6) doi: 10.1542/peds.2021-051772
123. Wu Q, Hu Y, Mo Z, et al. Development and Validation of a Deep Learning Model to Predict the Occurrence and Severity of Retinopathy of Prematurity. *JAMA Netw Open* 2022;5(6):e2217447. doi: 10.1001/jamanetworkopen.2022.17447 [published Online First: 20220601]
124. Frank E. Harrell J. Regression Modeling Strategies With Applications to Linear Models, Logistic and Ordinal Regression and Survival Analysis. 2nd ed. Switzerland: Springer International Publishing 2015.
125. Rodríguez G. Lecture Notes on Generalized Linear Models 2007 [Available from: https://grodri.github.io/glms/notes/ accessed February 27 2023 2023.
126. Cox DR. Regression Models and Life-Tables. *Journal of the Royal Statistical Society: Deries B (Methodological)* 1972;34:187-202. doi: https://doi.org/10.1111/j.2517-6161.1972.tb00899.x
127. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Stat Med* 2002;21(15):2175-97. doi: 10.1002/sim.1203
128. Holford TR. The Analysis of Rates and of Survivorship Using Log-Linear Models. *Biometrics* 1980;36(2):299-305.
129. Laird N, Olivier D. Covariance Analysis of Censored Survival Data Using Log-Linear Analysis Techniques. *Journal of the American Statistical Association* 1981;76(374):231-40.
130. Holford TR. Life Tables with Concomitant Information. *Biometrics* 1976;32(3):587-97.
131. Friedman M. Piecewise Exponential Models for Survival Data with Covariates. *Institute of Mathematical Statistics* 1982;10(1):101-13.
132. Whitehead J. Fitting Cox's Regression Model to Survival Data using GLIM. *Journal of the Royal Statistical Society* 1980;29(3):268-75.
133. Aitkin M, Clayton D. The Fitting of Exponential, Weibull and Extreme Value Distributions to Complex Censored Survival Data Using GLIM. *Journal of the Royal Statistical Society* 1980;29(2):156-63.
134. Frome E. The Analysis of Rates Using Poisson Regression Models. *Biometrics* 1983;39(3):665-74.
135. Hutchinson AK, Melia M, Yang MB, et al. Clinical Models and Algorithms for the Prediction of Retinopathy of Prematurity: A Report by the American Academy of Ophthalmology. *Ophthalmology* 2016;123(4):804-16. doi: 10.1016/j.ophtha.2015.11.003 [published Online First: 20160128]

136. Laws DE, Morton C, Weindling M, et al. Systemic effects of screening for retinopathy of prematurity. *Br J Ophthalmol* 1996;80(5):425-8. doi: 10.1136/bjo.80.5.425
137. Mitchell AJ, Green A, Jeffs DA, et al. Physiologic effects of retinopathy of prematurity screening examinations. *Adv Neonatal Care* 2011;11(4):291-7. doi: 10.1097/ANC.0b013e318225a332
138. Ranger M, Grunau RE. Early repetitive pain in preterm infants in relation to the developing brain. *Pain Manag* 2014;4(1):57-67. doi: 10.2217/pmt.13.61
139. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *BMC Med* 2015;13:1. doi: 10.1186/s12916-014-0241-z [published Online First: 20150106]
140. Steyerberg EW, Moons KG, van der Windt DA, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med* 2013;10(2):e1001381. doi: 10.1371/journal.pmed.1001381 [published Online First: 20130205]
141. Moons KGM, Wolff RF, Riley RD, et al. PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration. *Ann Intern Med* 2019;170(1):W1-W33. doi: 10.7326/M18-1377
142. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann Intern Med* 2019;170(1):51-58. doi: 10.7326/M18-1376
143. Binenbaum G. Algorithms for the prediction of retinopathy of prematurity based on postnatal weight gain. *Clin Perinatol* 2013;40(2):261-70. doi: 10.1016/j.clp.2013.02.004
144. Binenbaum G, Tomlinson LA, de Alba Campomanes AG, et al. Validation of the Postnatal Growth and Retinopathy of Prematurity Screening Criteria. *JAMA Ophthalmol* 2019 doi: 10.1001/jamaophthalmol.2019.4517 [published Online First: 2019/11/15]
145. Albertsson-Wikland K, Martensson A, Savendahl L, et al. Mortality Is Not Increased in Recombinant Human Growth Hormone-treated Patients When Adjusting for Birth Characteristics. *J Clin Endocrinol Metab* 2016;101(5):2149-59. doi: 10.1210/jc.2015-3951 [published Online First: 20160226]
146. FRAX - Fracture Risk Assessment Tool [Available from: https://www.sheffield.ac.uk/FRAX/index.aspx accessed April 11 2023 2023.
147. Carstensen B. Who needs the Cox model anyway? 2019 [Available from: http://bendixcarstensen.com/WntCma.pdf accessed January 13 2023 2023.
148. DW Hosmer SL. Applied Logistic Regression. 2nd ed. New York: John Wiley & Sons, Inc 2000.
149. Wu C, Lofqvist C, Smith LE, et al. Importance of early postnatal weight gain for normal retinal angiogenesis in very preterm infants: a multicenter study analyzing weight velocity deviations for the prediction of retinopathy of prematurity. *Arch Ophthalmol* 2012;130(8):992-9. doi: 10.1001/archophthalmol.2012.243
150. DIGIROP Prediction Models [Available from: http://www.digirop.com/index.html accessed April 11 2023 2023.
151. Obata S, Imamura T, Kakinoki M, et al. Systemic adverse events after screening of retinopathy of prematurity with mydriatic. *PLoS One* 2021;16(9):e0256878. doi: 10.1371/journal.pone.0256878 [published Online First: 20210909]
152. Agrawal Y, Patri S, Kalavakunta JK, et al. Retinopathy of prematurity screening leading to cardiopulmonary arrest: fatal complication of a benign procedure. *BMJ Case Rep* 2016;2016 doi: 10.1136/bcr-2016-216594 [published Online First: 20160728]

153. American Academy of Pediatrics Committee on F, Newborn, American Academy of Pediatrics Section on S, et al. Prevention and management of pain in the neonate: an update. *Pediatrics* 2006;118(5):2231-41. doi: 10.1542/peds.2006-2277
154. Tan Z, Isaacs M, Zhu Z, et al. Retinopathy of prematurity screening: A narrative review of current programs, teleophthalmology, and diagnostic support systems. *Saudi J Ophthalmol* 2022;36(3):283-95. doi: 10.4103/sjopt.sjopt_220_21 [published Online First: 20221014]
155. Kappen TH, van Loon K, Kappen MA, et al. Barriers and facilitators perceived by physicians when using prediction models in practice. *J Clin Epidemiol* 2016;70:136-45. doi: 10.1016/j.jclinepi.2015.09.008 [published Online First: 20150921]
156. Kantrowitz-Gordon I, Altman MR, Vandermause R. Prolonged Distress of Parents After Early Preterm Birth. *J Obstet Gynecol Neonatal Nurs* 2016;45(2):196-209. doi: 10.1016/j.jogn.2015.12.004 [published Online First: 20160123]
157. Huhtala M, Korja R, Lehtonen L, et al. Associations between parental psychological well-being and socio-emotional development in 5-year-old preterm children. *Early Hum Dev* 2014;90(3):119-24. doi: 10.1016/j.earlhumdev.2013.12.009 [published Online First: 20140110]
158. Schuetz Haemmerli N, Stoffel L, Schmitt KU, et al. Enhancing Parents' Well-Being after Preterm Birth-A Qualitative Evaluation of the "Transition to Home" Model of Care. *Int J Environ Res Public Health* 2022;19(7) doi: 10.3390/ijerph19074309 [published Online First: 20220404]
159. Gyllensten H, Humayun J, Sjobom U, et al. Costs associated with retinopathy of prematurity: a systematic review and meta-analysis. *BMJ Open* 2022;12(11):e057864. doi: 10.1136/bmjopen-2021-057864 [published Online First: 20221124]
160. Pivodic A, Hard AL, Lofqvist C, et al. Individual Risk Prediction for Sight-Threatening Retinopathy of Prematurity Using Birth Characteristics. *JAMA Ophthalmol* 2020;138(1):21-29. doi: 10.1001/jamaophthalmol.2019.4502
161. Pivodic A, Johansson H, Smith LEH, et al. Development and validation of a new clinical decision support tool to optimize screening for retinopathy of prematurity. *Br J Ophthalmol* 2021 doi: 10.1136/bjophthalmol-2020-318719 [published Online First: 20210512]
162. Pivodic A, L EHS, Hard AL, et al. Validation of DIGIROP models and decision support tool for prediction of treatment for retinopathy of prematurity on a contemporary Swedish cohort. *Br J Ophthalmol* 2022 doi: 10.1136/bjophthalmol-2021-320738 [published Online First: 20220311]
163. Good WV, Hardy RJ, Dobson V, et al. The incidence and course of retinopathy of prematurity: findings from the early treatment for retinopathy of prematurity study. *Pediatrics* 2005;116(1):15-23. doi: 10.1542/peds.2004-1413
164. Palmer EA, Flynn JT, Hardy RJ, et al. Incidence and early course of retinopathy of prematurity. The Cryotherapy for Retinopathy of Prematurity Cooperative Group. *Ophthalmology* 1991;98(11):1628-40. doi: 10.1016/s0161-6420(91)32074-8
165. Quinn GE, Johnson L, Abbasi S. Onset of retinopathy of prematurity as related to postnatal and postconceptional age. *Br J Ophthalmol* 1992;76(5):284-8. doi: 10.1136/bjo.76.5.284
166. Chiang MF, Jiang L, Gelman R, et al. Interexpert agreement of plus disease diagnosis in retinopathy of prematurity. *Arch Ophthalmol* 2007;125(7):875-80. doi: 10.1001/archopht.125.7.875

167. Slidsborg C, Forman JL, Fielder AR, et al. Experts do not agree when to treat retinopathy of prematurity based on plus disease. *Br J Ophthalmol* 2012;96(4):549-53. doi: 10.1136/bjophthalmol-2011-300573 [published Online First: 20111215]
168. Bolon-Canedo V, Ataer-Cansizoglu E, Erdogmus D, et al. Dealing with inter-expert variability in retinopathy of prematurity: A machine learning approach. *Comput Methods Programs Biomed* 2015;122(1):1-15. doi: 10.1016/j.cmpb.2015.06.004 [published Online First: 20150616]
169. IBM - What are neural networks? [Available from: https://www.ibm.com/topics/neural-networks accessed April 11 2023 2023.
170. Kappen TH, van Klei WA, van Wolfswinkel L, et al. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. *Diagn Progn Res* 2018;2:11. doi: 10.1186/s41512-018-0033-6 [published Online First: 20180612]
171. Bauer MS, Kirchner J. Implementation science: What is it and why should I care? *Psychiatry Res* 2020;283:112376. doi: 10.1016/j.psychres.2019.04.025 [published Online First: 20190423]
172. Lloyd Minor MR. Discovering Precision Health - Predict, Prevent, and Cure to Advance Health and Well-being. Oxford: John Wiley & Sons, Ltd 2020.

APPENDICES

APPENDIX 1 HAZARD FUNCTIONS FROM DIGIROP-BIRTH

DIGIROP-BIRTH 1.0 PREDICTION MODEL

The parameter estimates obtained for the final DIGIROP-Birth 1.0 by using in-house developed SAS-macro for the extended Poisson regression.

TABLE 6. DIGIROP-BIRTH 1.0 PREDICTION MODEL. SE = *standard error; BW = birth weight; PND = parenteral nutrition duration; PNA = postnatal age; w = weeks; GA = gestational age; BWSDS = birth weight standard deviation score; SDS = standard deviation score; INT = interaction.*

X	Variable	β	Estimate (SE)	p-value
X_0	Intercept	β_0	-20.1666 (4.9219)	<0.0001
X_1	PNA 0 to 8w, per 1w increase	β_1	1.7331 (0.6129)	0.0050
X_2	PNA >8 to 12w, per 1w increase	β_2	0.3618 (0.0992)	0.0003
X_3	PNA >12w, per 1w increase	β_3	-0.3788 (0.0857)	<0.0001
X_4	GA 24-27w, per 1w increase (centered at 28w)	β_4	-0.8210 (0.3353)	0.014
X_5	GA >27w, per 1w increase (centered at 28w)	β_5	0.7266 (0.7302)	0.32
X_6	Sex, 1=*boys*, 2=*girls*	β_6	-0.9385 (0.3054)	0.0021
X_7	BWSDS ≤-1 SDS, per 1SDS increase	β_7	0.1521 (0.2656)	0.57
X_8	BWSDS >-1 SDS, per 1SDS increase	β_8	-1.0401 (0.4710)	0.027
X_9	INT: PNA × GA 24-27w	β_9	0.0227 (0.0230)	0.32
X_{10}	INT: PNA × GA >27w	β_{10}	-0.1360 (0.0627)	0.030
X_{11}	INT: Sex × GA	β_{11}	-0.2505 (0.1066)	0.019
X_{12}	INT: PNA × BWSDS ≤-1SDS	β_{12}	-0.0371 (0.0199)	0.062
X_{13}	INT: PNA × BWSDS >-1SDS	β_{13}	0.0728 (0.0349)	0.037

DIGIROP-BIRTH 1.0 HAZARD CALCULATION

The hazard function for a given individual i is calculated as

$$\lambda_i = exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \beta_8 X_{8i} + \beta_9 X_{9i} + \beta_{10} X_{10i} + \beta_{11} X_{11i} + \beta_{12} X_{12i} + \beta_{13} X_{13i})$$

where β_0 to β_{13} are the parameter estimates from ***Table 6*** above, and X as per below

$X_0 = 1$

$X_1 = min(t, 8)$

$X_2 = min(max(t, 8), 12 - 8)$

$X_3 = max(x - 12, 0)$

$X_4 = min(GAC, -1)$

$X_5 = max(GAC - (-1), 0)$

$X_6 = SEX$

$X_7 = min(BWSDS, -1)$

$X_8 = max(BWSDS - (-1), 0)$

$X_9 = t\ x\ min(GAC, -1)$

$X_{10} = t\ x\ max(GAC - (-1), 0)$

$X_{11} = SEX\ x\ GAC$

$X_{12} = t\ x\ min(BWSDS, -1)$

$X_{13} = t\ x\ max(BWSDS - (-1), 0)$

where $SEX = 1$ for *boys*, $SEX = 2$ for *girls*, and $GAC = GA - 28$

Calculating the hazard function per small intervals for $t = [0,20]$ and applying those in the numerical integration, the survival function $S(t)$ and the cumulative distribution function $F(t)$ are obtained. $F(t)$ is providing the risk estimates for ROP treatment that is of interest for the project.

$$F(t) = 1 - S(t) = 1 - exp\left(-\int_0^{20} \lambda(x)dx\right)$$

DIGIROP-BIRTH 2.0 PREDICTION MODEL

The parameter estimates obtained for the final DIGIROP-Birth 2.0 by using in-house developed SAS-macro for the extended Poisson regression.

TABLE 7. DIGIROP-BIRTH 2.0 PREDICTION MODEL. SE = *standard error; BW = birth weight; PND = parenteral nutrition duration; PNA = postnatal age; w = weeks; GA = gestational age; BW = birth weight; PND = parenteral nutrition duration INT = interaction.*

X	Variable	β	Estimate (SE)	p-value
X_0	Intercept	β_0	-17.2284 (2.8183)	<0.0001
X_1	PNA 0 to 8w, per 1w increase	β_1	1.6218 (0.3523)	<0.0001
X_2	PNA >8 to 12w, per 1w increase	β_2	0.4897 (0.0483)	<0.0001
X_3	PNA >12w, per 1w increase	β_3	-0.3505 (0.0234)	<0.0001
X_4	GA 24-27w, per 1w increase (centered at 28w)	β_4	0.0446 (0.1245)	0.72
X_5	GA >27w, per 1w increase (centered at 28w)	β_5	1.8942 (0.5803)	0.0011
X_6	Sex, 0=*boys*, 1=*girls*	β_6	-0.9740 (0.2767)	0.0004
X_7	BW per 100g increase	β_7	-0.3331 (0.0538)	<0.0001
X_8	PN ≥14d *vs.* <14d	β_8	0.4744 (0.1509)	0.0017
X_9	PN Unknown *vs.* <14d	β_9	0.1963 (0.1662)	0.24
X_{10}	INT: PNA × GA >27w	β_{10}	-0.1579 (0.0505)	0.0018
X_{11}	INT: Sex × PN ≥14d *vs.* <14d	β_{11}	0.4317 (0.2120)	0.042
X_{12}	INT: BW × GA	β_{12}	-0.0709 (0.0184)	0.0001
X_{13}	INT: Sex × GA 24-27w	β_{13}	-0.1528 (0.0740)	0.039

DIGIROP-BIRTH 2.0 HAZARD CALCULATION

The hazard function for a given individual i is calculated as

$$\lambda_i = exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \beta_8 X_{8i} + \beta_9 X_{9i} + \beta_{10} X_{10i} + \beta_{11} X_{11i} + \beta_{12} X_{12i} + \beta_{13} X_{13i})$$

where β_0 to β_{13} are the parameter estimates from ***Table 7*** above, and X as per below

$X_0 = 1$

$X_1 = min(t, 8)$

$X_2 = min(max(t, 8), 12 - 8)$

$X_3 = max(x - 12,0)$

$X_4 = min(GAC, -1)$

$X_5 = max(GAC - (-1), 0)$

$X_6 = SEX$

$X_7 = BW100$

$X_8 = PND1$

$X_9 = PND2$

$X_{10} = t\ x\ max(GAC - (-1), 0)$

$X_{11} = SEX\ x\ PND1$

$X_{12} = \text{BW100}\ x\ GA$

$X_{13} = SEX\ x\ min(GAC, -1)$

where $SEX = 0$ for *boys*, $SEX = 1$ for *girls*, $GAC = GA - 28$, $PND1 = 1$ for *≥14 days*, $PND1 = 0$ otherwise, $PND2 = 1$ for *unknown*, $PND2 = 0$ otherwise, $BW100 = BW/100$

The survival function $S(t)$ and the cumulative distribution function $F(t)$ are obtained in the same way as for DIGIROP-Birth 1.0.

DIGIROP-BIRTH 2.0 R CODE

```{r}
#home directory for the files
setwd("*mydirectory*")
#read register data
SNQDEV <- read.table("./*myfile*.csv",T,",")
```

Required packages

```{r, message=FALSE}
library(Epi)
library(popEpi)
library(survival)
library(mgcv)
```

Poisson regression model by Holford

```{r}
#create a lexis object
SNQDEVx <- Lexis(exit = list(tfe=EVLtreatWK), exit.status = factor(evltreat, labels=c("0","1")),
data = SNQDEV)
#split data in small intervals
SNQDEVx.s <- splitMulti(SNQDEVx, tfe=c(0,sort(unique(SNQDEVx$EVLtreatWK))))
```

```{r}
#create re-parametrized variables for continuous variables
SNQDEVx.s$timeL <- pmin(SNQDEVx.s$tfe,8)
SNQDEVx.s$timeM <- pmin(pmax(SNQDEVx.s$tfe-8,0),12-8)
SNQDEVx.s$timeU <- pmax(SNQDEVx.s$tfe-12,0)
SNQDEVx.s$gacL <- pmin(SNQDEVx.s$gac,-1)
SNQDEVx.s$gacU <- pmax(SNQDEVx.s$gac+1,0)
```

```{r}
#Final DIGIROP-Birth Poisson model with time modeled with piecewise linear function with
selected break points corresponding to the model run in SAS
m.pois <- glm(cbind(lex.Xst=="1",lex.dur) ~ timeL + timeM + timeU + gacL + gacU + sex01 +
bw100 + PNGE14_1 + PNGE14_2 + tfe:gacU + sex01:PNGE14_1 + gac:bw100 + gacL:sex01,
family=poisreg, data=SNQDEVx.s)
summary(m.pois)
```

The parameter estimates obtained for the DIGIROP-Birth 2.0 by using the *glm* function in R for the extended Poisson regression. Small differences in parameter estimates are expected for the reasons of estimation procedures applied and convergence limits used. For 11139 included infants in *Paper IV*, the difference in risk estimates for R–SAS method, given the respective parameter estimates, was mean 0.0015, median 0.000045, 5th percentile -0.000039, and 95th percentile 0.0045.

```
Call:
glm(formula = cbind(lex.Xst == "1", lex.dur) ~ timeL + timeM +
    timeU + gacL + gacU + sex01 + bw100 + PNGE14_1 + PNGE14_2 +
    tfe:gacU + sex01:PNGE14_1 + gac:bw100 + gacL:sex01, family = poisreg,
    data = SNQDEVx.s)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.3432  -0.0205  -0.0062  -0.0014   4.8958

Coefficients:
               Estimate Std. Error z value Pr(>|z|)
(Intercept)    -11.51405    1.81881  -6.331 2.44e-10 ***
timeL            0.87845    0.22488   3.906 9.37e-05 ***
timeM            0.55526    0.04912  11.304  < 2e-16 ***
timeU           -0.34964    0.02300 -15.201  < 2e-16 ***
gacL             0.04082    0.12338   0.331 0.740772
gacU             1.77958    0.37174   4.787 1.69e-06 ***
sex01           -0.97704    0.27709  -3.526 0.000422 ***
bw100           -0.33042    0.05364  -6.160 7.27e-10 ***
PNGE14_1         0.47193    0.15079   3.130 0.001750 **
PNGE14_2         0.19364    0.16618   1.165 0.243921
gacU:tfe        -0.15079    0.02933  -5.141 2.73e-07 ***
sex01:PNGE14_1   0.43339    0.21199   2.044 0.040909 *
bw100:gac       -0.07001    0.01820  -3.848 0.000119 ***
gacL:sex01      -0.15338    0.07405  -2.071 0.038350 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 7104.5  on 885697  degrees of freedom
Residual deviance: 5142.8  on 885684  degrees of freedom
AIC: 6064.8

Number of Fisher Scoring iterations: 13
```

APPENDIX 2 ONLINE APPLICATION INFORMATION

DIGIROP-BIRTH 1.0 PREDICTION MODEL

DIGIROP-Birth
www.digirop.com

RETINOPATHY OF PREMATURITY PREDICTION MODELS
WWW.ROP.GU.SE
The Sahlgrenska Center for Pediatric Ophthalmology Research

DIGIROP-Birth is an early-estimating risk model for Type I ROP (ROP treatment) based only on the infant's gestational age (GA) at birth, birth weight, and sex, for infants born at 24+0 to 30+6 weeks of gestation.[1]

Instructions

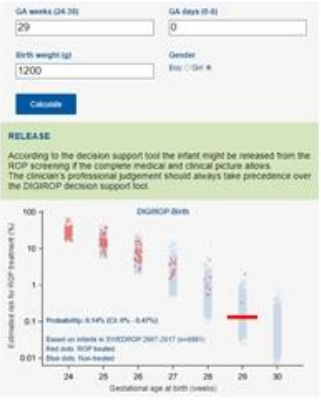
- Enter GA weeks (possible values 24-30)
- Enter GA days (possible values 0-6)
- Enter Birth weight in grams
- Tick if Boy or Girl
- Press Calculate

Results

- Estimated risk (%) is provided with 95% CI
- Clinical decision support tool provides either SCREEN or RELEASE based on the estimated risk.
- Figure shows GA-specific estimated risks from the development population, blue dots represent non-treated and red dots ROP treated infants. Red line is the current infant's value.

www.digirop.com

Example Infant



Caution: DIGIROP models are not reliable for infants diagnosed with severe congenital malformations/syndromes, hydrocephalus, and those that have performed intestinal surgery (e.g. for necrotizing enterocolitis). These infants should be screened as per routine standards.

THIS MODEL IS DEVELOPED BY *THE SAHLGRENSKA CENTER FOR PEDIATRIC OPHTHALMOLOGY RESEARCH*. IT IS BASED ON ~300 EVENTS FROM ~7000 INFANTS SCREENED FOR ROP IN SWEDEN YEARS 2007-2018 (SWEDROP REGISTRY), AND VALIDATED ON EXTERNAL DATA FROM SWEDEN, GERMANY, AND USA.

Any questions, please contact: carola.pfeiffer-mosesson@gu.se

[1] Pivodic A, Hård AL, Löfqvist C, Smith LEH, Wu C, Bründer MC, Lagrèze WA, Stahl A, Holmström G, Albertsson-Wikland K, Johansson H, Nilsson S, Hellström A. Individual Risk Prediction for Sight-Threatening Retinopathy of Prematurity Using Birth Characteristics. JAMA Ophthalmol. 2020 Jan 1;138(1):21-29. doi: 10.1001/jamaophthalmol.2019.4502.

DIGIROP-SCREEN 1.0 PREDICTION MODEL

DIGIROP-Screen
www.digirop.com

RETINOPATHY OF PREMATURITY PREDICTION MODELS
WWW.ROP.GU.SE
The Sahlgrenska Center for Pediatric Ophthalmology Research

DIGIROP-Screen is a risk model for Type I ROP (ROP treatment) based on the infant's DIGIROP-Birth risks (GA, birth weight, sex), and the time of the first ROP diagnosis, aimed for use along the screening for infants born at 24+0 to 30+6 weeks of gestation.[1]

Instructions
- Enter GA weeks (possible values 24-30)
- Enter GA days (possible values 0-6)
- Enter Birth weight in grams
- Tick if Boy or Girl
- Enter Birth date (yyyy-mm-dd)
- Enter Date for prediction (yyyy-mm-dd)
- Tick if ROP diagnosed including date for first diagnosis (yyyy-mm-dd), or not
- Press Calculate

Results
- Estimated risk (%) is provided with 95% CI
- Clinical decision support tool provides either SCREEN or RELEASE based on the estimated risk.
- Figure shows GA-specific estimated risks from the development population, blue dots represent non-treated and red dots ROP treated infants. Red line is the current infant's value.

www.digirop.com
Example Infant

Caution: DIGIROP models are not reliable for infants diagnosed with severe congenital malformations/syndromes, hydrocephalus, and those that have performed intestinal surgery (e.g. for necrotizing enterocolitis). These infants should be screened as per routine standards.

THIS MODEL IS DEVELOPED BY THE SAHLGRENSKA CENTER FOR PEDIATRIC OPHTHALMOLOGY RESEARCH. IT IS BASED ON ~300 EVENTS FROM ~7000 INFANTS SCREENED FOR ROP IN SWEDEN YEARS 2007-2018 (SWEDROP REGISTRY), AND VALIDATED ON EXTERNAL DATA FROM SWEDEN, GERMANY, AND USA.
Any questions, please contact: carola.pfeiffer-mosesson@gu.se

[1] Pivodic A, Johansson H, Smith LEH, Hård AL, Löfqvist C, Yoder BA, Hartnett ME, Wu C, Bründer MC, Lagrèze WA, Stahl A, Al-Hawasi A, Larsson E, Lundgren P, Gränse L, Sunnqvist B, Tornqvist K, Wallin A, Holmström G, Albertsson-Wikland K, Nilsson S, Hellström A. Development and validation of a new clinical decision support tool to optimize screening for retinopathy of prematurity. Br J Ophthalmol. 2021 May 12:bjophthalmol-2020-318719. doi: 10.1136/bjophthalmol-2020-318719.